RAJA ALMUKKAHAL

PRE-CALCULUS

CK-12 Foundation

Pre-Calculus

Say Thanks to the Authors
Click http://www.ck12.org/saythanks
(No sign in required)

Almukkahal Fiori Fortgang Landers Vigil

Contents

1 Analyzing Conic Sections **1**

1.1 Introduction to Conic Sections . . . 1
1.2 Circles and Ellipses . . . 9
1.3 Parabolas . . . 28
1.4 Hyperbolas . . . 44
1.5 General Algebraic Forms . . . 57

2 Analyzing Functions **63**

2.1 Identifying Functions . . . 63
2.2 Minimums and Maximums . . . 73
2.3 Increasing and decreasing functions . . . 79
2.4 End Behavior of Functions . . . 89
2.5 Function families . . . 96
2.6 Transformations of functions . . . 106
2.7 Operations and Compositions . . . 118
2.8 Headline text . . . 124
2.9 Functions and Mathematical Models . . . 125

3 Analyzing Polynomial and Rational Functions **137**

3.1 Quadratic Functions . . . 137
3.2 Polynomial Functions . . . 152
3.3 Rational Functions . . . 167
3.4 Analyzing Rational Functions . . . 183
3.5 Polynomial and Rational Inequalities . . . 191
3.6 Finding Real Zeros of Polynomial Functions . . . 200
3.7 Approximating Real Zeros of Polynomial Functions . . . 213
3.8 The Fundamental Theorem of Algebra . . . 218

4 Analyzing Exponential and Logarithmic Functions **223**

4.1 Inverse Functions . . . 223
4.2 Exponential Functions . . . 232
4.3 Logarithmic Functions . . . 248
4.4 Properties of Logarithms . . . 258
4.5 Exponential and Logarithmic Models and Equations . . . 265
4.6 Compound Interest . . . 275
4.7 Growth and Decay . . . 287
4.8 Applications . . . 295

5 Polar Equations and Complex Numbers **300**

5.1 Polar Coordinates . . . 300
5.2 Polar-Cartesian Transformations . . . 304
5.3 Systems of Polar Equations . . . 315
5.4 Imaginary and Complex Numbers . . . 325
5.5 Operations on Complex Numbers . . . 329
5.6 Trigonometric Form of Complex Numbers . . . 334
5.7 Product and Quotient Theorems . . . 338
5.8 Powers and Roots of Complex Numbers . . . 343

6 Vectors **351**

6.1 Vectors in a Plane . . . 351
6.2 Vectors in Space . . . 362
6.3 Dot Products . . . 371
6.4 Cross Products . . . 381
6.5 Planes in Space . . . 389
6.6 Vector Direction . . . 396
6.7 Vector Equations . . . 402
6.8 Applications of Vector Analysis . . . 407

7 Introduction to Calculus **414**

7.1 Limits (An Intuitive Approach) . . . 414
7.2 Computing Limits . . . 423
7.3 Tangent Lines and Rates of Change . . . 429

8 Sequences, Series, and Mathematical Induction **435**

8.1 Recursive and Explicit Formulas 435
8.2 Summation Notation 442
8.3 Mathematical Induction 447
8.4 Mathematical Induction, Factors, and Inequalities 454
8.5 Geometric Series 462
8.6 The Binomial Theorem 468

Chapter 1

Analyzing Conic Sections

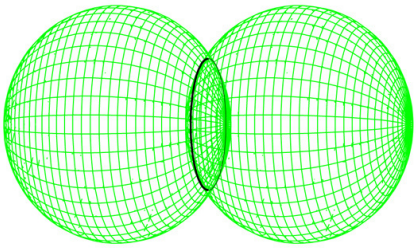
1.1 Introduction to Conic Sections

Learning Objectives

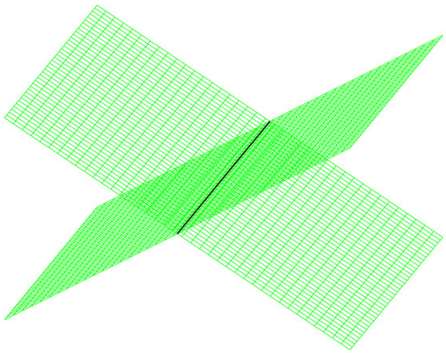
- Consider the results when two simple mathematical objects are intersected.
- Be comfortable working with an infinite two-sided cone.
- Know the basic types of figures that result from intersecting a plane and a cone.
- Know some of the history of the study of conic sections.

Introduction: Intersections of Figures

Some of the best mathematical shapes come from intersecting two other important shapes. Two spheres intersect to form a circle:

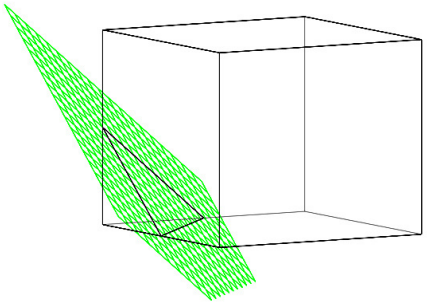


Two planes intersect to form a line:

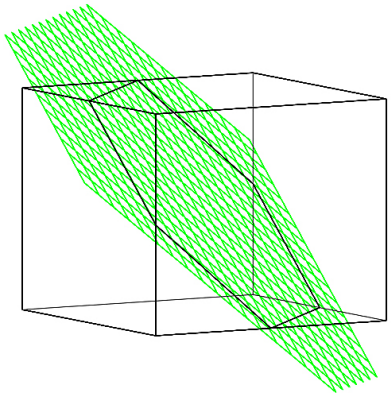


While simple and beautiful, for these two examples of intersections there isn't much else to investigate. For the spheres, no matter how we put them together, their intersection is always either nothing, a point (when they just touch), a circle of various sizes, or a sphere if they happen to be exactly the same size and coincide. Once we've exhausted this list, the inquiry is over. Planes are even simpler: the intersection of two distinct planes is either nothing (if the planes are parallel) or a line.

But some intersections yield more complex results. For instance a plane can intersect with a cube in numerous ways. Below a plane intersects a cube to form an equilateral triangle.



Here a plane intersects a cube and forms a regular hexagon.



Review Questions

1. Describe all the types of shapes that can be produced by the intersection of a plane and a cube.
2. What is the side-length of the regular hexagon that is produced in the above diagram when the cube

has side-length 1?

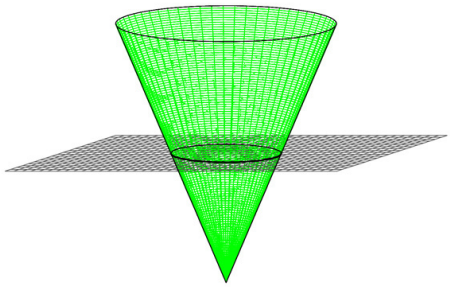
Review Answers

1. Square, rectangle, pentagon, others.
2. $length = \frac{\sqrt{2}}{2}$

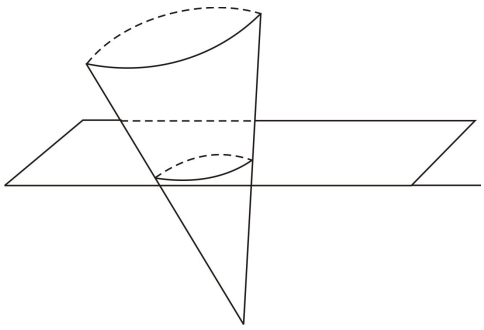
Intersections with Cones

One class of intersections is of particular interest: The intersection of a plane and a cone. These intersections are called **conic sections** and the first person known to have studied them extensively is the Ancient Greek mathematician Menaechmus in the 3rd Century B. C. E. Part of his interest in the conic sections came from his work on a classic Greek problem called "doubling the cube," and we will describe this problem and Menaechmus' approach that uses conic sections in section three. The intersections of the cone and the plane are so rich that the resulting shapes have continued to be of interest and generate new ideas from Menaechmus' time until the present.

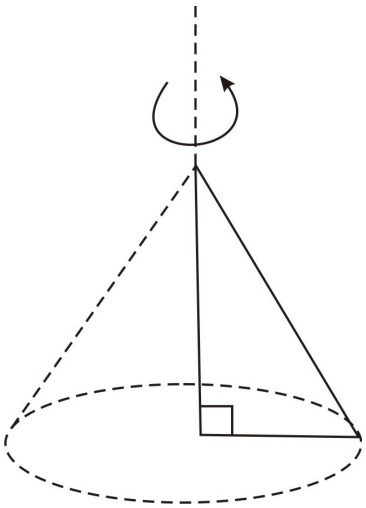
Before we really delve into what we mean by a plane and a cone, we can look at an intuitive example. Suppose by a cone we just mean an ice-cream cone. And by a plane we mean a piece of paper. Well, if you sliced through an upright ice-cream cone with a horizontal piece of paper you would find that the two objects intersect at a circle.



That's very nice. But unlike the intersection of two spheres, which also resulted in a circle, that's not *all* we get. If we tilt the paper (or the ice cream cone) things start to get tricky.

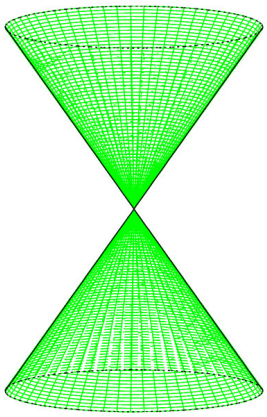


First things first, we better make sure we know what we mean by "plane" and "cone". Let's use the simplest definitions possible. So by "plane", we mean the infinitely thin flat geometric object that extends forever in all directions. Even though infinity is a tricky concept, this plane is in some sense simpler than one that ends arbitrarily. There is no boundary to think about with the infinite plane. And what do we mean by a cone? An ice-cream cone is a good start. In fact, it's very similar to how the ancient Greeks defined a cone, as a right triangle rotated about one if its legs.



Like we do with the infinite geometric plane, we want to idealize this object a bit too. As with the plane, to avoid having to deal with a boundary, let's suppose it continues infinitely in the direction of its open end.

But the Greek mathematician Apollonius noticed that it helps even more to have it go to infinity in the other direction. This way, a cone can be thought of as an infinite collection of lines, and since geometric lines go on forever in both directions, a cone also extends to infinity in both directions. Here is a picture of what we will call a **cone** in this chapter (remember it extends to infinity in both directions).



A cone can be formally defined as a three-dimensional collection of lines, all forming an equal angle with a central line or axis. In the above picture, the central line is vertical.

Review Questions

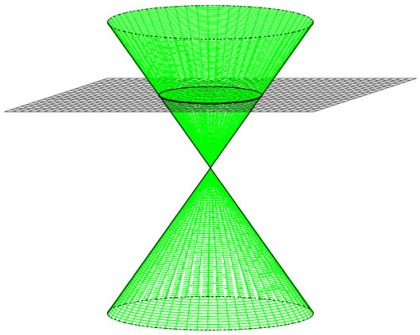
3. What other mathematical objects can be generated by a collection of lines?

Review Answers

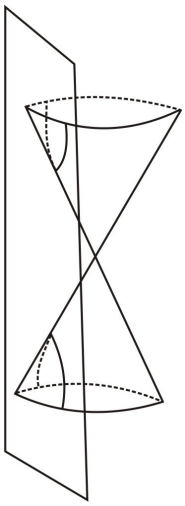
3. Cylinder (infinite), plane, three-sided infinite pyramids, others.

Intersections of an Infinite Cone and Plane

Like with the ice cream cone and the piece of paper, some intersections of this infinite cone with an infinite plane yield a circle.

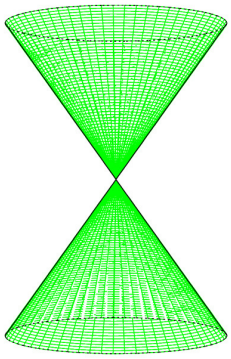


Other intersections yield something a bit less tidy. In the picture below, the plane is parallel to the axis of the cone.



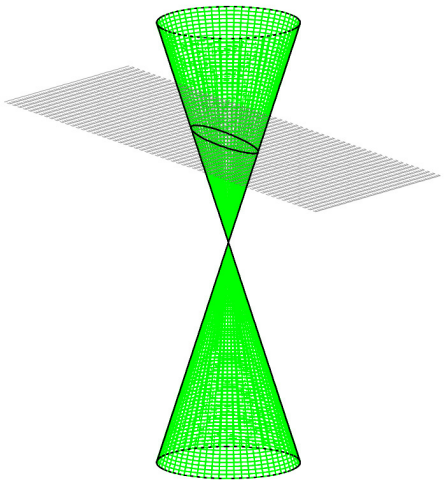
But it turns out that the set of intersections of a cone and a plane forms a beautiful, mathematically consistent, set of shapes that have interesting properties. So in this chapter we embark on a study of these intersections called *conic sections*.

Note: Although we will not prove it here, it doesn't matter if the cone is asymmetrical or tilted to one side (also called "oblique"). If you include "tilted cones" the same conic sections result. So, for simplicity's sake let's stick with non-tilted, or "right" cones and focus on what happens

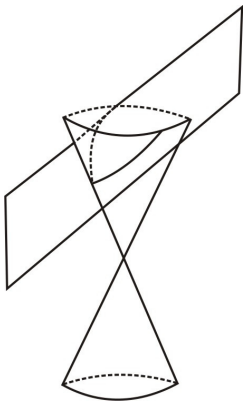


Let's begin by doing a rough tally of the kinds of shapes it seems we can generate by slicing planes through our cone. First of all, we have the **circle**, as we discovered with the ice cream cone above. A circle is formed when the plane is perpendicular to the line in the middle of the cone.

If we tilt the plane a little, but not so much that it intersects both cones, we get something more oval shaped. This is called an **ellipse**. Later you will learn many of the fascinating properties of the ellipse.



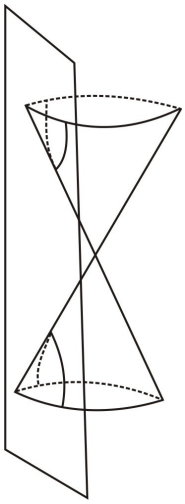
If really tilt the plane more so that it only intersects one side of the cone, but we get a big infinite "U".



This shape is called a **parabola** and like the ellipse it has a number of surprising properties.

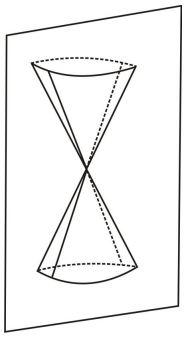
Then if we tilt it even further, we intersect both sides of the cone and get two big "U's" going in opposite

directions.



This pair of objects is called a hyperbola, and, like the parabola and ellipse, the hyperbola has a number of interesting properties that we will discuss.

Finally, in a few cases we get what are called "degenerate" conics. For instance, if we line up a vertical plane with the vertex of the cone, we get two crossing lines.



Review Questions

4. Are there any other types of intersections between a plane and a cone besides the ones illustrated above?
5. Is it possible for a plane to miss a cone entirely?

Review Answers

4. Yes, the plane could meet the vertex, resulting in an intersection of a single point.
5. It is not possible for a plane to miss a cone entirely since both of the objects extend infinitely.

Applications and Importance

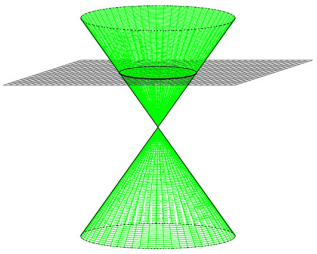
The intersections of cones and planes produce an interesting set of shapes, which we will study in the upcoming sections. We will also see numerous applications of these shapes to the physical world. Why is it that the intersection of a cone and a plane would produce so many applications? We don't seem to see

a lot of cones or planes in our daily life. But at closer inspection they are everywhere. Point sources of light, approximated by such sources as a flashlight or the sun, shine in cone-shaped array, so for instance the image of a flashlight against a slanted wall is an ellipse.

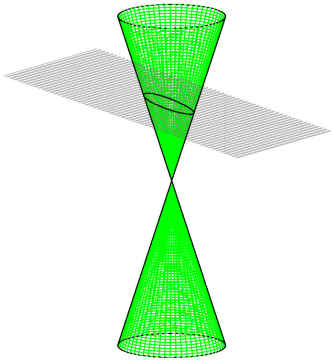
The geometric properties of conics, such as the focal property of the ellipse, turn out to have many physical ramifications, such as the design of telescopes. And, when we inspect the algebraic representations of conic sections, we will see that there are similarities with the law of gravity, which in turn has ramifications for planetary motion.

Lesson Summary

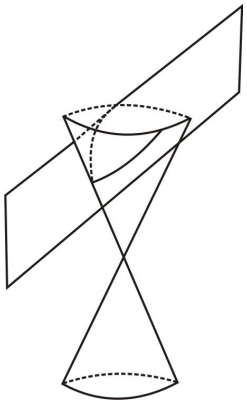
In summary, here are some of the ways that a plane can intersect a cone.



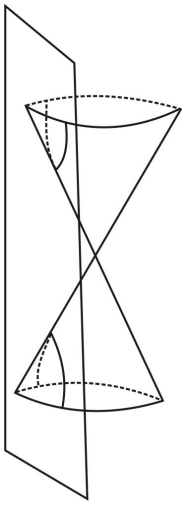
A circle



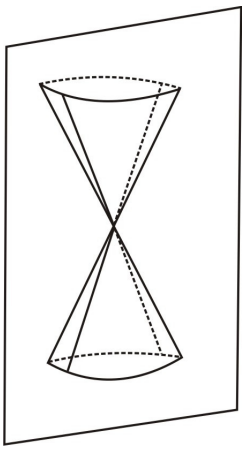
an "oval" called the ellipse



A big infinite "U" called a parabola.



Two big, infinite "U"s called a hyperbola.



Strange "degenerate" shapes like two crossing lines, as well as other examples you may have found in Review Question 1.

This array of shapes has a surprising amount of mathematical coherence, as well as a large number of interesting properties.

Vocabulary

Cone A three-dimensional collection of lines, all forming an equal angle with a central line or axis.

Conic section The points of intersection between a cone and a plane.

1.2 Circles and Ellipses

Learning Objectives

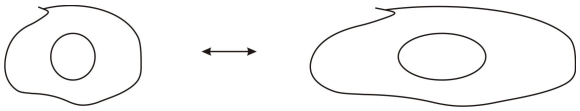
- Understand the difference between an "oval" and an ellipse.
- Recognize and work with equations for ellipses.
- Derive the focal property of ellipses.

- Understand the equivalence of different definitions of ellipses.
- Reconstruct Dandelin's sphere construction.
- Know some of the different ways people have approached ellipses throughout history.
- Understand some of the important applications of ellipses.

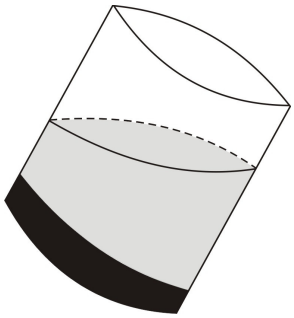
Introduction

Let's begin with the first class of shapes discussed in the last section. When the plane makes a finite intersection with one side of the cone, we get either a circle or the "oval-shaped" object illustrated in the previous section. It turns out that this is no ordinary oval, but something called an **ellipse**, a shape with special properties.

Like parallelograms, or any other shape with lots of interesting properties, ellipses can be defined by some of these properties, and then the other properties necessarily follow from the definition. For example, a parallelogram is typically defined as a quadrilateral with each pair of opposite pairs of sides parallel. Once you define it this way, it follows that the opposite sides must also be equal in length, and that the diagonals must bisect each other. Well, if you instead started by defining a parallelogram by one of these other properties, for instance opposite sides having equal lengths, then you would end up with the same class of shapes. The same thing happens for ellipses. One way to define an ellipse is as a "stretched out circle". It's the shape you would get if you sketched a circle on a deflated balloon and then stretched out the balloon evenly in two opposite directions:



It's also the shape of the surface of water that results when you tilt a round glass:



Or an ellipse could be thought of as the shape of a circle drawn on a piece of paper when it is viewed at an angle.

Equations of Ellipses

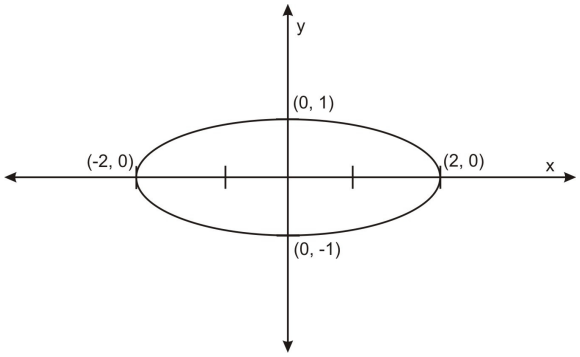
This "stretching" can be represented algebraically. For simplicity, take the circle of radius 1 centered at the origin (0,0). The distance formula tells us that this is the set of points (x, y) that is a distance 1 unit away from the origin.

$$\begin{aligned} D &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \\ 1 &= \sqrt{(x-0)^2 + (y-0)^2} \\ 1 &= x^2 + y^2 \end{aligned}$$

This equation, $x^2 + y^2 = 1$, can be altered to stretch the circle in the horizontal (i.e. $x-$axis) direction by dividing the x variable by a constant $a > 1$,

$$\left(\frac{x}{a}\right)^2 + y^2 = 1$$

Why does this stretch the circle horizontally? Well, the effect of dividing x by a is that for each $y-$value in an ordered pair (x, y) that satisfies the original equation, the corresponding x value must be multiplied by a in order for the pair to make a solution to the altered equation. So solutions (x, y) of the circle are in one-to-one correspondence with solutions (ax, y) of the altered equation, hence stretching the corresponding graph to the left and right by a factor of a. For example, here is the graph of $\left(\frac{x}{2}\right)^2 + y^2 = 1$:



Generalizing the equation by allowing a stretch in the vertical direction, we get the following.

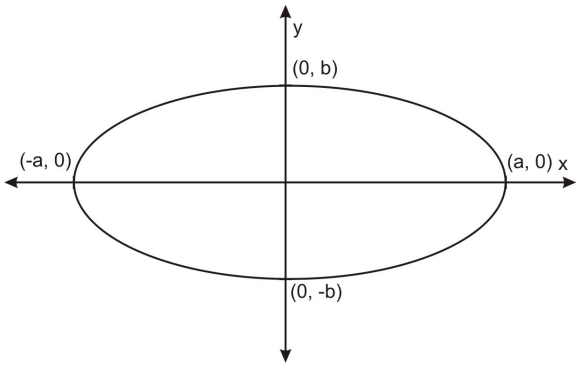
$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

The factor a stretches the circle in the horizontal direction and the factor b stretches the circle in the vertical direction. If $a = b$, this is just a circle. When $a \neq b$, this equation represents an ellipse. The ellipse is stretched in the horizontal direction if $b < a$ and it is stretched in the vertical direction if $a < b$. Often the above equation is written as follows.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

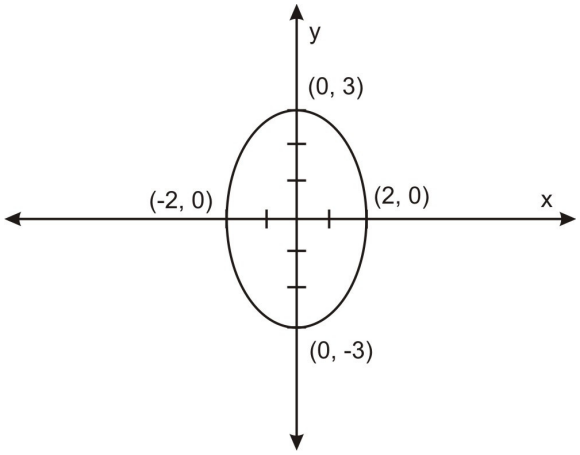
This is called the standard form of the equation of an ellipse, assuming that the ellipse is centered at (0,0).

To sketch a graph of an ellipse with the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, start by plotting the four axes intercepts, which are easy to find by plugging in 0 for x and then for y. Then sketch the ellipse freehand, or with a graphing program or calculator.



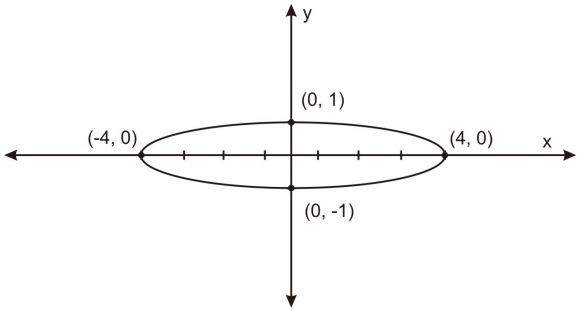
Example 1 *Sketch the graph of* $\frac{x^2}{4} + \frac{y^2}{9} = 1$.

Solution: This equation can be rewritten as $\frac{x^2}{2^2} + \frac{y^2}{3^2} = 1$. After setting $x = 0$ and $y = 0$ to find the $y-$ and $x-$intercepts, (0,3), (0,-3), (2,0), and (-2,0), we sketch the ellipse about these points:



Example 2 *Sketch the graph of* $\frac{x^2}{16} + y^2 = 1$.

Solution: This can be rewritten as $\frac{x^2}{4^2} + \frac{y^2}{1^2} = 1$. After finding the intercepts and sketching the graph, we have:

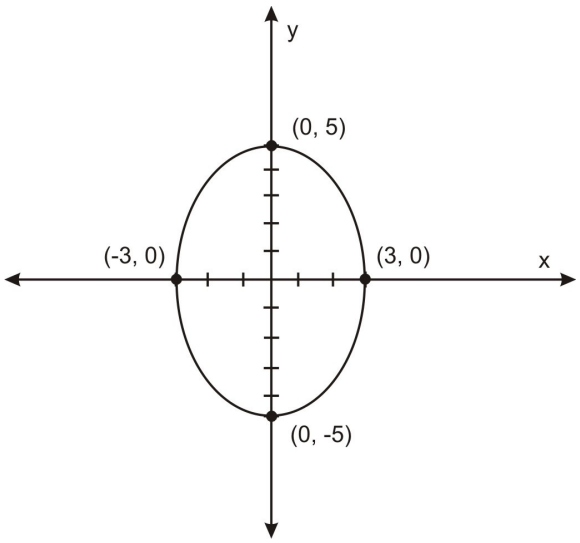


The segment spanning the long direction of the ellipse is called the **major axis**, and the segment spanning the short direction of the ellipse is called the **minor axis**. So in the last example the major axis is the segment from (-4,0) to (4,0) and the minor axis is the segment from (0,-1) to (0,1).

The major and minor axes are examples of what are sometimes called **reference lines**. Apollonius, the Ancient Greek mathematician who wrote an early treatise on conics, used these and other reference lines to orient conic sections. Though the Greeks did not use a coordinate plane to discuss geometry, these reference lines offer a framing perspective that is similar to the Cartesian plane that we use today. Apollonius' way of framing conics with reference lines was the closest mathematics came to the system of coordinate geometry that you know so well until Descartes' and Fermat's systematic work in the seventeenth century.

Example 3 *Not all equations for ellipses start off in the standard form above. For example,* $25x^2+9y^2 = 225$ *is an ellipse. Put it in the proper form and graph it.*

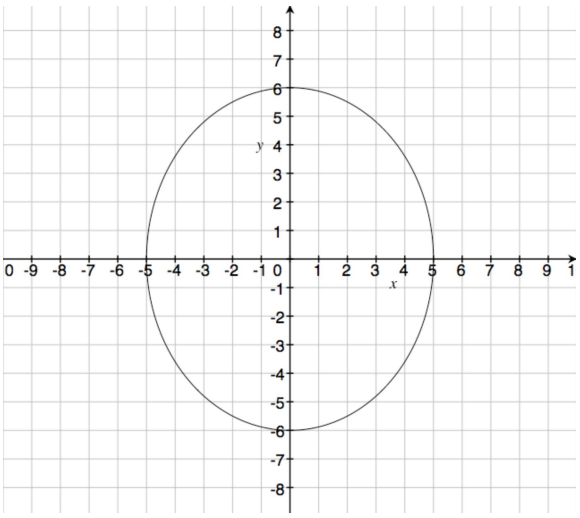
Solution: First, divide both sides by 225, to get: $\frac{x^2}{9} + \frac{y^2}{25} = 1$, or $\frac{x^2}{3^2} + \frac{y^2}{5^2} = 1$. Finding the intercepts and graphing, we have:



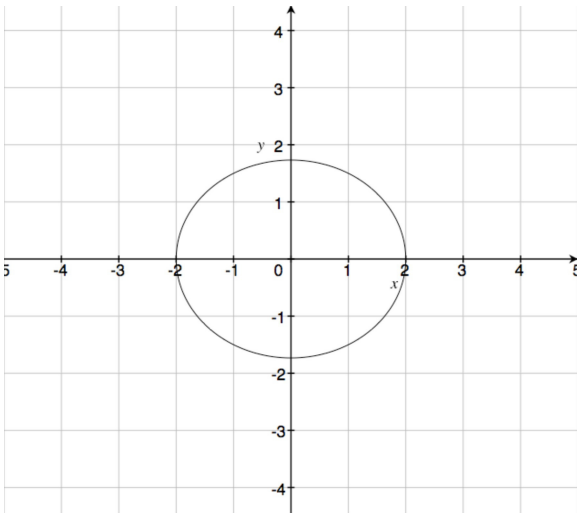
Review Questions

1. It was mentioned above that when a round glass of water is tilted, the surface of the water is an ellipse. Using our working definition of an ellipse as "stretched out circle", explain why you think the water takes this shape.
2. Sketch the following ellipse: $36x^2 + 25y^2 = 900$
3. Now try sketching this ellipse where the numbers don't turn out to be so neat: $3x^2 + 4y^2 = 12$

Review Answers

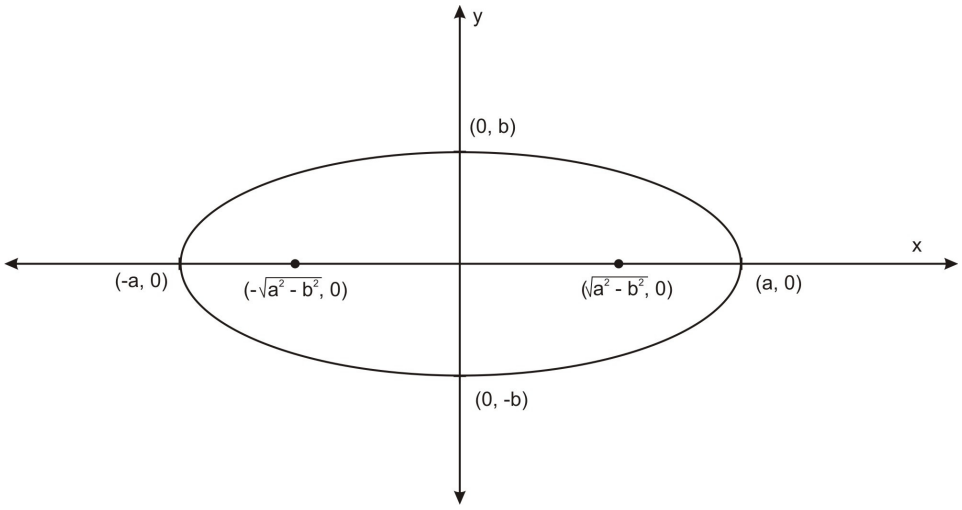
1. Answers may vary, but should explain why the shape that results stretches a circle in one direction because the width of the glass is constant.
2. 

3.

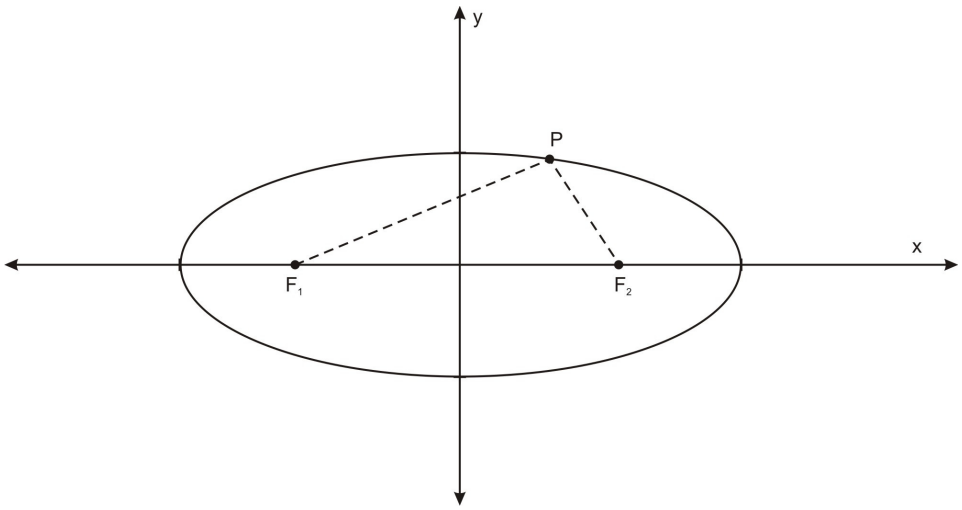


The Focal Property

In every ellipse there are two special points called the **foci** (foci is plural, **focus** is singular), which lie inside the ellipse and which can be used to define the shape. For an ellipse centered at (0,0) that is wider than it is tall, its major axis is horizontal and its foci are at $\left(\sqrt{a^2-b^2}, 0\right)$ and $\left(-\sqrt{a^2-b^2}, 0\right)$.



What is the significance of these points? The ellipse has a geometric property relating to these points that is similar to a circle's relationship with its center. Remember a circle can be thought of as the set of points in a plane that are a certain distance from the center point. In fact, that is typically the definition of a circle. Well, the foci act like the center except that there are two of them. An ellipse is the set of points where the *sum* of the distance between each point on the ellipse and each of the two foci is a constant number. In the diagram below, for any point P on the ellipse, $F_1P + F_2P = d$, where F_1 and F_2 are the foci and d is a constant.



Technology

This definition using the sum of the focal distance gives us a great way to draw ellipses. Sure, you can just graph them on your calculator. But why not utilize a simpler technology that does the job just as well? As you know, a circle could be drawn by fixing a string to a piece of paper, tying the other end to a pencil, and then drawing the curve that keeps the string taut.

Similarly, an ellipse can be drawn by taking a string that is longer than the distance between two points, fixing the two ends of the string to the two points. Then drawing all points that can be drawn when the string is taut and the pencil is touching it. In the diagram above, the dotted line represents the string of fixed length, which is attached at the foci F_1 and F_2. The string is looped around the pencil at P, and then the pencil is moved, keeping the string taut, drawing all the points whose sum of distances between F_1 and F_2 is the length of the string.

Review Questions

4. Use string and tacks to draw ellipses that
   (a) are nearly circles
   (b) are very different than circles

For each of these, what can you say about how the distance between the foci and the length of the string compare?

Review Answers

4. Drawings may vary. For ellipses that are nearly circles, the distance between the foci is small compared to the length of string.

The foci can also be used to measure how far an ellipse is "stretched" from a circle. The symbol ε stands for the **eccentricity** of an ellipse, and it is defined by the distance between the foci divided by the length of the major axis, or $\frac{\sqrt{a^2-b^2}}{a}$ for horizontally oriented ellipses and $\frac{\sqrt{b^2-a^2}}{b}$ for vertically oriented ellipses. Since a circle is an ellipse where $a = b$, circles have an eccentricity of 0.

Review Questions

5. What is the full range of the eccentricity of an ellipse? What does it look like near the extremes of this range?

Review Answers

5. The interval of possible values is $\varepsilon \in [0,1)$. At $\varepsilon = 0$, the ellipse is a circle; as the eccentricity approaches 1 it becomes more and more elongated.

Turning the Definition of Ellipses on its Head

Often, this focal property is not though of as a *property* of ellipses, but rather a *defining feature*. To see that these are equivalent, we have some to work to do. Let's start by proving that stretched out circles actually have this focal property. We want to prove that for a "stretched out circle" defined by the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where a and b are two positive numbers, the sum of the distance between every point on the stretched circle and the two foci $\left(\sqrt{a^2-b^2}, 0\right)$ and $\left(-\sqrt{a^2-b^2}, 0\right)$ is the same. So we need to prove that for every point on the shape defined by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, the sum of the distances to $\left(\sqrt{a^2-b^2}, 0\right)$ and $\left(-\sqrt{a^2-b^2}, 0\right)$ is the same number.

Proof: Suppose a point (x,y) is on the curve defined by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Then we can solve for y in the equation and express the point in terms of x:

$$\begin{aligned} \frac{y^2}{b^2} &= 1 - \frac{x^2}{a^2} \\ y^2 &= \frac{a^2b^2 - b^2x^2}{a^2} \\ y &= \sqrt{\frac{a^2b^2 - b^2x^2}{a^2}} \end{aligned}$$

Now, using the distance formula to compute the sum of the distance between the pairs of points: $\left(x, \sqrt{\frac{a^2b^2-b^2x^2}{a^2}}\right)$ and $\left(\sqrt{a^2-b^2}, 0\right)$; and, $\left(x, \sqrt{\frac{a^2b^2-b^2x^2}{a^2}}\right)$ and $\left(-\sqrt{a^2-b^2}, 0\right)$ We have:

$$\sqrt{\left(x - \sqrt{a^2-b^2}\right)^2 + \frac{a^2b^2 - b^2x^2}{a^2}} + \sqrt{\left(x + \sqrt{a^2-b^2}\right)^2 + \frac{a^2b^2 - b^2x^2}{a^2}}$$

This algebraic equation looks daunting, but simplifying the expressions inside these square roots results in a surprising result. This becomes

$$\sqrt{\frac{a^2\left(x-\sqrt{a^2-b^2}\right)^2+a^2b^2-b^2x^2}{a^2}}+\sqrt{\frac{a^2\left(x+\sqrt{a^2-b^2}\right)^2+a^2b^2-b^2x^2}{a^2}}$$
$$=\frac{1}{a}\left(\sqrt{a^2\left(x-\sqrt{a^2-b^2}\right)^2+a^2b^2-b^2x^2}+\sqrt{a^2\left(x+\sqrt{a^2-b^2}\right)^2+a^2b^2-b^2x^2}\right)$$
$$=\frac{1}{a}\left(\sqrt{a^2x^2-2a^2x\sqrt{a^2-b^2}+a^4-a^2b^2+a^2b^2-b^2x^2}+\sqrt{a^2x^2-2a^2x\sqrt{a^2-b^2}+a^4-a^2b^2+a^2b^2-b^2x^2}\right)$$
$$=\frac{1}{a}\left(\sqrt{a^2x^2-2a^2x\sqrt{a^2-b^2}+a^4-b^2x^2}+\sqrt{a^2x^2-2a^2x\sqrt{a^2-b^2}+a^4-b^2x^2}\right)$$
$$=\frac{1}{a}\left(\sqrt{\left(a^2-x\sqrt{a^2-b^2}\right)^2}+\sqrt{\left(a^2+x\sqrt{a^2-b^2}\right)^2}\right)$$
$$=\frac{1}{a}\left(a^2-x\sqrt{a^2-b^2}+a^2+x\sqrt{a^2-b^2}\right)$$
$$=\frac{1}{a}(2a^2)$$
$$=2a$$

What a miraculous collapse! One of the gnarliest algebraic expressions that I have ever encountered turned into the simple expression $2a$. Most importantly, this reduced expression is merely a constant—it doesn't depend on x or y. This is exactly what we were hoping for. *The sum of the distances between every point on the ellipse and the two foci is always* $2a$. If you every find yourself needing to remember what this sum is, the number $2a$ can be remembered most easily by computing it from an easy point such as one of the shape's $x-$intercepts, say $(a,0)$.

Review Questions

6. Compute the distance from the $x-$intercept $(a,0)$ and the two foci and show that it is in fact $2a$.
7. What is the sum of the distances to the foci of the points on a vertically-oriented ellipse?

Review Answers

6. The distance between the $x-$intercept $(a,0)$ and $\left(\sqrt{a^2-b^2},0\right)$ is: $a-\sqrt{a^2-b^2}$. The distance between the $x-$intercept $(a,0)$ and $\left(-\sqrt{a^2-b^2},0\right)$ is: $a+\sqrt{a^2-b^2}$.
   Together these add to: $a-\sqrt{a^2-b^2}+a+\sqrt{a^2-b^2}=2a$
7. $2b$

Defining an Ellipse by Focal Distance

So all this computing simply means that "stretched-out circles"—what we've been calling ellipses—satisfy the focal property. What would be great is if we could *define* ellipses by the focal property. This would be a nice generalization of the way we define circles.

Recall that circles are defined as the set of points in a plane that are a constant distance from a center point. Analogously, ellipses could be defined as a set of points in a plane for which the sum of the distances

to two focus points is a constant. In other words, this definition would yield the exact same set of shapes as the "stretched out circle" definition that we started with.

Before proceeding we need to make sure of one thing. We've already proved that stretched out circles satisfy the focal property, but how do we know that any shape satisfying the focal property is in turn a stretched out circle? Well, the calculation above can simply be read backwards. In other words, suppose you have a set of points that satisfy the focal property, that each point whose sum of the distances to the points $(f,0)$ and $(-f,0)$ is a fixed distance d. Now note that for any two positive numbers d and f with $2f < d$, there exist positive numbers a and b such that $d = 2a$ and $f = \sqrt{a^2 - b^2}$. (this fact is a bit subtle and is part of an exercise below.) So now we have a shape where every point (x,y) has a sum of distances to the points $\left(\sqrt{a^2 - b^2}, 0\right)$ and $\left(-\sqrt{a^2 - b^2}, 0\right)$ which equals $2a$. Using these expressions, the algebraic steps of the proof above can simply be read backwards.

Review Questions

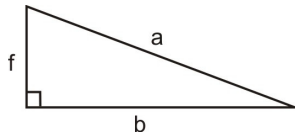
8. Explain why for any two positive numbers d and f with $2f < d$, there exist positive numbers a and b such that $d = 2a$ and $f = \sqrt{a^2 - b^2}$.
9. We just told you that the above proof could be read backwards. But you need to be careful when following algebraic steps backwards, especially ones involving squares or square roots.
   (a) For example, what happens when you follow this argument backwards?

   $$x = -2$$
   $$x^2 = 4$$

   (b) Write a convincing argument that it is okay to follow the steps backwards in the above proof that every stretched circle has the focal property.

Review Answers

8. Set $a = \frac{d}{2}$. Then we need to show that for f satisfying $2f < d$, there exists a number b such that $f = \sqrt{a^2 - b^2}$. Since $2f < d, 2f < 2a$ by the definition of a (using the assumption $d > 0$). So $f < a$. We can find b geometrically. Since $f < a$, there is a right triangle with one leg having length f and hypotenuse a. Call the other leg of the triangle b.



   Then the Pythagorean Theorem tells us that $f^2 + b^2 = a^2$, or equivalently $f = \sqrt{a^2 - b^2}$. So a and b satisfy the necessary requirements.
9. (a) Taking the square root of both sides of $x^2 = 4$ yields two solutions, $x = \pm 2$, instead of the one value we already know ($x = -2$). The problem is that the operation of squaring a number is not a one-to-one function. Both $(-2)^2$ and 2^2 yield the same number. So some information is lost during this step, and it cannot be perfectly "undone", like other algebraic maneuvers.
   (b) The student's reason should include the fact no information is lost (through squaring both sides or other operations) in any of these steps, so that each step is completely reversible.

Equation of an Ellipse Not Centered at the Origin

All the ellipses we've looked at so far are centered around the origin (0,0). To find an equation for ellipses centered around another point, say (h,k), simply replace x with $x - h$ and y with $y - k$. This will shift all

the points of the ellipse to the right h units (or left if $h < 0$) and to up k units (or down if $k < 0$). So the general form for a horizontally- or vertically-oriented ellipse is:

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

It is centered about the point (h, k). If $b < a$, the ellipse is horizontally oriented and has foci $\left(h + \sqrt{a^2 - b^2}, k\right)$ and $\left(h - \sqrt{a^2 - b^2}, k\right)$ on its horizontal major axis. If $a < b$, it is vertically oriented and has foci $\left(h, k + \sqrt{a^2 - b^2}\right)$ and $\left(h, k - \sqrt{a^2 - b^2}\right)$ on its vertical major axis.

Example 4

Graph the equation $4x^2 + 8x + 9y^2 - 36y + 4 = 0$.

Solution : We need to get the equation into the form of general equation above. The first step is to group all the x terms and y terms, factor our the leading coefficients of x^2 and y^2, and move the constants to the other side of the equation:

$$4(x^2 + 2x) + 9(y^2 - 4y) = -4$$

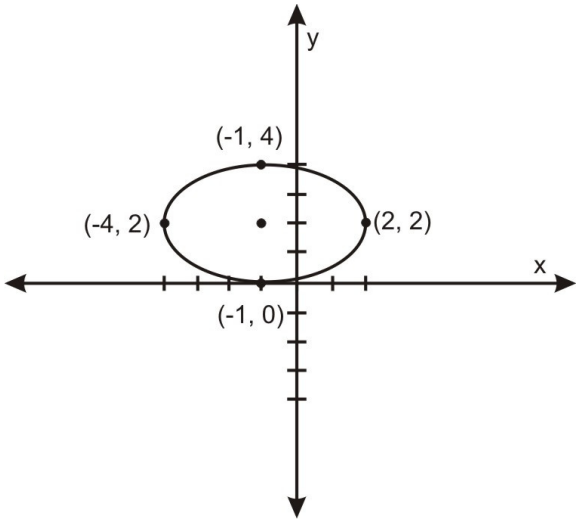
Now, we "complete the square" by adding the appropriate terms to the x expressions and the y expressions to make a perfect square. (See http://authors.ck12.org/wiki/index.php/Algebra_I-Chapter-10#Solving_Quadratic_Equations_by_Completing_the_Square for more on completing the square.)

$$4(x^2 + 2x + 1) + 9(y^2 - 4y + 4) = -4 + 4 + 36$$

Now we factor and divide by the coefficients to get:

$$\frac{(x+1)^2}{9} + \frac{(y-2)^2}{4} = 1$$

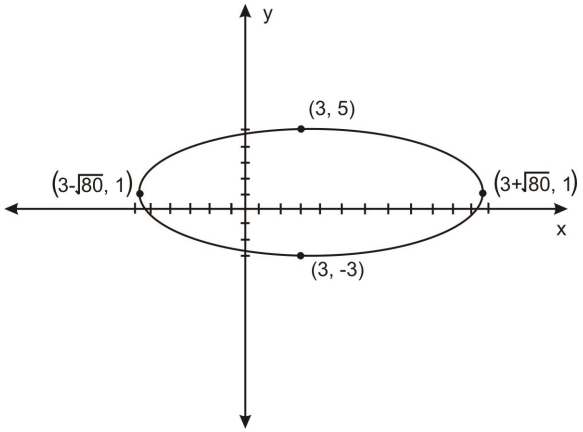
And there we have it. Once it's in this form, we see this is an ellipse is centered around the point (-1,2), it has a horizontal major axis of length 3 and a vertical minor axis of length 2, and from this we can make a sketch of the ellipse:

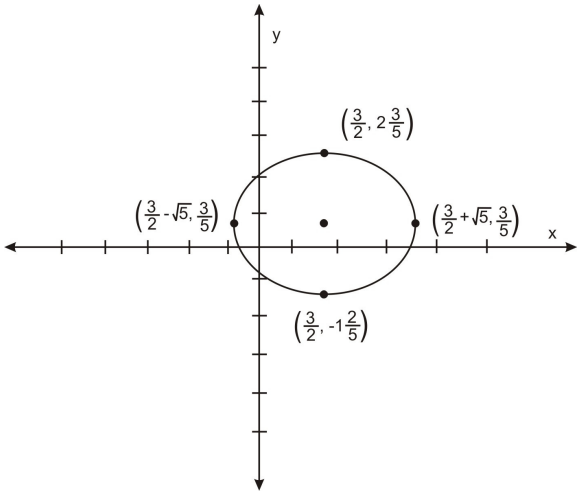


Review Questions

10. Explain why subtracting h from the $x-$term and k from the $y-$term in the equation for an ellipse shifter the ellipse h horizontally and k vertically.
11. Graph this ellipse. $x^2 - 6x + 5y^2 - 10y - 66 = 0$
12. Now try this one that doesn't have such nice numbers!
$$16x^2 - 48x + 125y^2 + 150y + 61 = 0$$
13. Now try this one. $3x^2 - 12x + 5y^2 + 10y - 3 = 0$. What goes wrong? Explain what you think the graph of this equation might look like.
14. What about this one. $5x^2 - 15x - 2y^2 + 8y - 50 = 0$? What goes wrong here? Explain what you think the graph of this equation might look like.

Review Answers

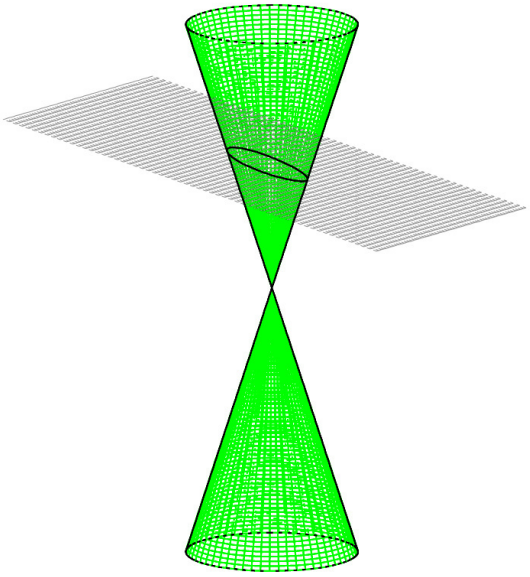
10. If (x, y) is a solution to $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ then $(x+h, y+k)$ is a solution to $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$. This produces a graph that is shifted horizontally by h and vertically by k.
11. 

12. 

13. After completing the square, we have the sum of positive numbers equaling a negative number. This is an impossibility, so the equation has no solutions.
14. After completing the square, the x term and the y term are opposite signs. If you plot some points you will see that the graph has two disconnected sections. This class of conic sections will be discussed in the next section.

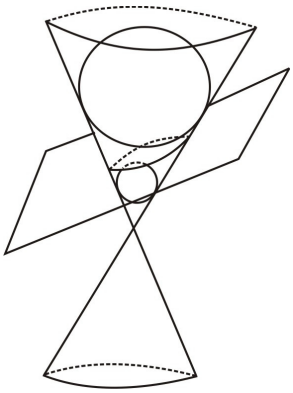
Difference Between an Ellipse and an Oval or Proving the Sliced Cone Definition of an Ellipse

There is still one critical step missing in our exploration of ellipses. We showed that "stretched-out circles" satisfy the focal property, and that any shape satisfying this property is in fact a "stretched-out circle". So these are actually the same class of shapes, and they are called ellipses. But not any oval-shaped curve is an ellipse. Draw a random oval and you're not likely to be able to find two points that satisfy the focal property. In particular, when we cut a cone with a tilted plane, how do we know that the oval-shaped curve that results is a "stretched out circle" satisfying the focal property?

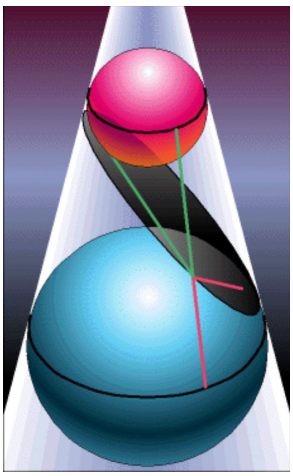


Amazingly, the Ancient Greeks had an argument for this fact over two millennia ago. While it is impressive that this problem was solved so long ago, the argument itself involves an intricate construction that isn't as illuminating as a more modern the one I'm going to show you instead. Most mathematicians prefer to use a more modern argument that is simply stunning in its simplicity. This modern argument isn't fancy—the Greeks had all the tools they need to understand it—they just didn't happen to think of it. It wasn't until 1822 that the French mathematician Germinal Dandelin thought of this very clever construction. Dandelin found a way to find the foci and prove the focal property in one fell swoop. Here's what he said.

Take the conic section in question. Then choose a sphere that is just the right size so that when it's dropped into the conic, it touches the intersecting plane, as well as being snug against the cone on all sides. If you prefer, you can think of the sphere as a perfectly round balloon that is blown up until it "just fits" inside the cone, still touching the plane. Then do the same on the *other* side of the plane. fter we've drawn both of these spheres we have this picture:

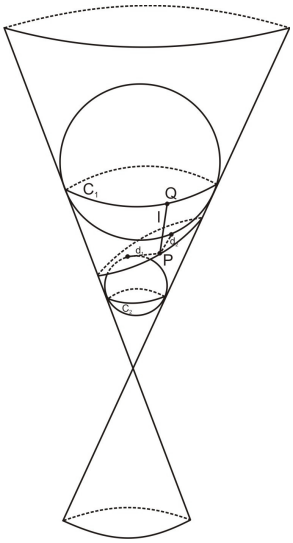


or

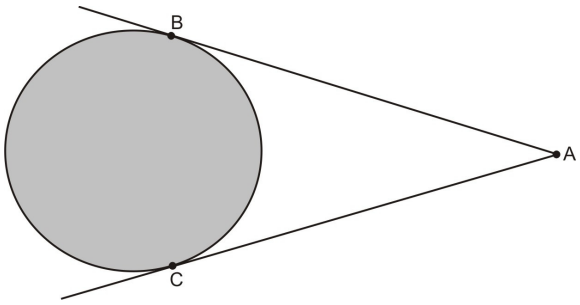


These spheres are often called "Dandelin spheres", named after their discoverer. It turns out that not only is our shape an ellipse (which, like all ellipses satisfies the focal property). But these spheres touch the ellipse exactly at the two foci. To see this, consider this geometric argument.

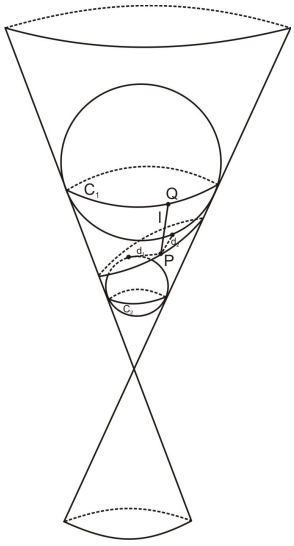
The first thing to notice is that the circles C_1 and C_2 shown on the diagram below, where each sphere lies snug against the cone, lie in parallel planes to one another. In particular, each line passing through these circles and the vertex of the cone, such as the line l drawn below, cuts off equal segments between the two circles. Let's call d the shortest distance along the cone between circles C_1 and C_2. This can also be thought of as the shortest distance between C_1 and C_2 that passes through the vertex of the cone.



The next thing to remember is a property of tangents to spheres that you may have learned in geometry. If two segments are drawn between a point and a sphere, and if the line containing each segment is tangent to the sphere, then the two segments are equal. In the diagram below, $AB = AC$. (This follows from the fact that tangents are perpendicular to the radii of a sphere and that two congruent triangles are formed in this configuration. See this description for more about this property.)



Now consider the point P on the ellipse drawn below. Let $\overline{QR}$ be the segment of length d between C_1 and C_2 that passes through P. The distances between the two foci are marked d_1 and d_2. But $d_1 = RP$ and $d_2 = PQ$ by the property of tangents to spheres discussed above. So $d_1 + d_2 = RP + PQ = QR = d$. And this sum will always equal d, no matter what point P on the ellipse is chosen. So this proves the focal property of ellipses: that the sum of the distances between any point on the ellipse and the two foci is constant.



Review Questions

15. What do the Dandelin spheres look like in the case of a circle?
16. What is the area of an ellipse with the equation $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$? (Hint: use a geometric argument starting with the area of a circle.)
17. What is the perimeter of an ellipse with the equation $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$?

Review Answers

15. The Dandelin spheres for a circle lie directly above one another, and both touch the circle at the center point.
16. The area of an ellipse is $ab\pi$. To see why this is true, start with a circle of radius 1, which has an area of π. Then imagine an approximation with rectangles of the circle. Then stretch the rectangles by a factor of a in the $x-$direction and by a factor of b in the $y-$direction to obtain an approximation of the ellipse. This makes the rectangles a times wider and b times taller, giving an area that is ab multiplied by the area of the approximation of the circle. Since this is true of any approximation of

the circle, the area of the ellipse must be $ab\pi$.

17. This is actually a much more difficult question than the previous one. You're on your own! Even the great Indian mathematician Ramanujan could only come up with an approximation: $p \approx \pi\left[3(a+b) - \sqrt{(3a+b)(a+3b)}\right]$.

Applications

The number of places ellipses appear in the natural world is immense. Consider, for instance, how often the simplest kind of ellipse, the circle, appears in your life. You see circles emanating as waves when you throw a rock into a pond. You see circular pupils when you look at a set of eyes. You see what is roughly a circle as the image of the sun or moon in the sky. The path of an object swung around on a string is a circle. When any of these circles are viewed at an angle you see an elongated circle, or an ellipse. So it is important to keep in mind that the discussion that follows covers only a few of the instances of ellipses in our daily lives.

Planetary Motion

When a planet orbits the sun (or when any object orbits any other), it takes an elliptical path and the sun lies at one of the two foci of the ellipse. Johannes Kepler first proposed this at the beginning of the seventeenth century as one of three laws of planetary motions, after analyzing observational data of Tycho Brahe. His law is accurate enough to produce modern computations which are still used to predict the motion of artificial satellites. A century later, Newton's law of gravity offers an explanation of why this law might be true.

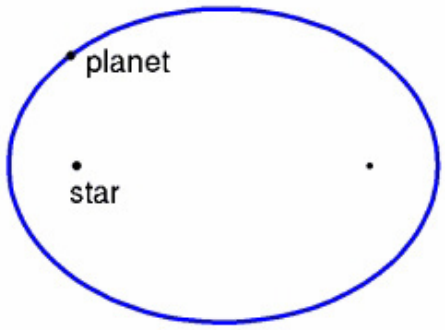


Figure 1.1

Review Questions

18. Though planets take an elliptical path around the sun, these ellipses often have a very low eccentricity, meaning they are close to being circles. The diagram above exaggerates the elliptical shape of a planet's orbit. The Earth's orbit has an eccentricity of 0.0167. Its minimum distance from the sun is 146 million km. What is its maximum distance from the sun? If the sun's diameter is 1.4 million kilometers. Do both foci of the Earth's orbit lie within the sun? Recall that the eccentricity of an ellipse is $\varepsilon = \frac{\sqrt{a^2-b^2}}{a}$.
19. While the elliptical paths of planets are ellipses that are closely approximated by circles, comets and asteroids often have orbits that are ellipses with very high eccentricity. Halley's comet has an

eccentricity of 0.967, and comes within 54.6 million miles of the sun at its closest point, or "perihelion". What is the furthest point it reaches from the sun?

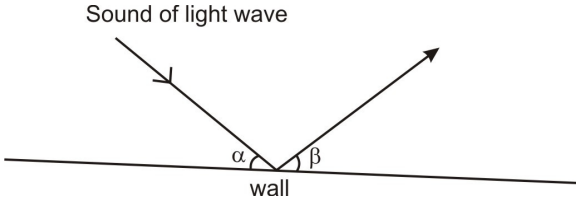
Review Answers

18. Assume that the orbit of the sun is an ellipse centered at (0,0). Then we can use the distance from the origin to the focus $\sqrt{a^2-b^2}$ to set up the equations $146+146+2\sqrt{a^2-b^2}=2a$ and $0.167=\frac{\sqrt{a^2-b^2}}{a}$. Solving we get $a=175.270, b=175.245$, and the distance from (0,0) to the foci, $c=2.927$ (all units are in millions of km). Finally the maximum distance from the earth to the sun is approximately 152 million km. From Kepler's law, we know one of the foci of its orbit is at the center of the sun. The other foci is $2(2.927)=5.854$ million kilometers away, so it is outside the sun (but not by very far!).
19. ~ 3.25 billion miles.

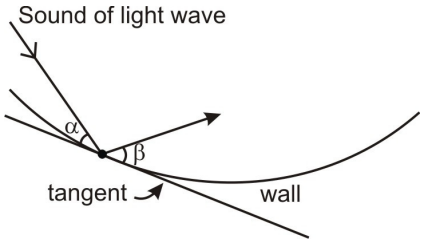
Echo Rooms

The National Statuary Hall in the United States Capital Building is an example of an ellipse-shaped room, sometimes called an "echo room", which provide an interesting application to a property of ellipses. If a person whispers very quietly at one of the foci, the sound echoes in a way such that a person at the other focus can often hear them very clearly. Rumor has it that John Quincy Adams took advantage of this property to eavesdrop on conversations in this room.

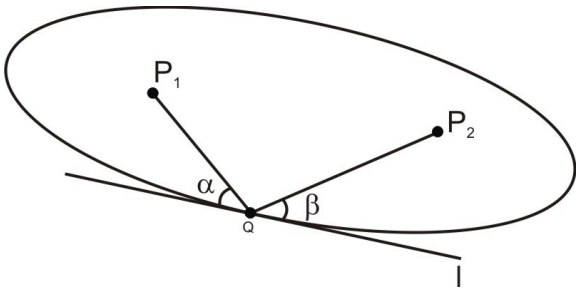
The property of ellipses that makes echo rooms work is called the "optical property." So why echoes, if this is an optical property? Well, light rays and sound waves bounce around in similar ways. In particular, they both bounce off walls at equal angles. In the diagram below, $\alpha=\beta$.



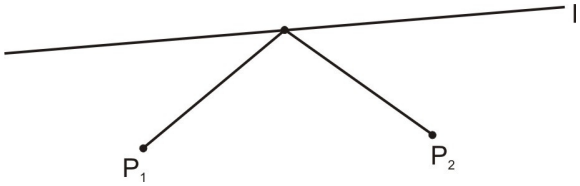
For a curved wall, they bounce at equal angles to the tangent line at that point:



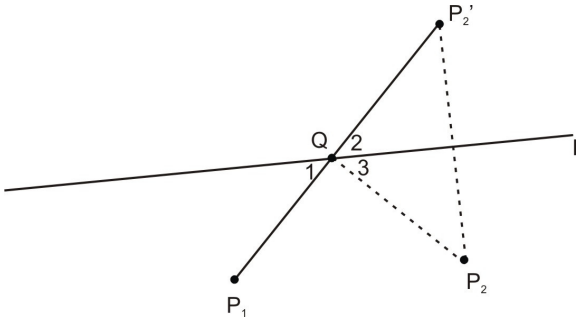
So the "optical property" of ellipses is that lines between a point on the ellipse and the two foci form equal angles to the tangent at that point, or in other words, whispers coming from one foci bounce directly to the other foci. In the diagram below, for each Q on the ellipse, $\angle\alpha \cong \angle\beta$.



This seems reasonable, given the symmetry of the ellipse, but how do we know it is true? First, let's prove an important property that is a bit more general. Suppose you have two points, P_1 and P_2, and a line l, and you are interested in the shortest possible path between two P_1 and P_2, that intersects with l.



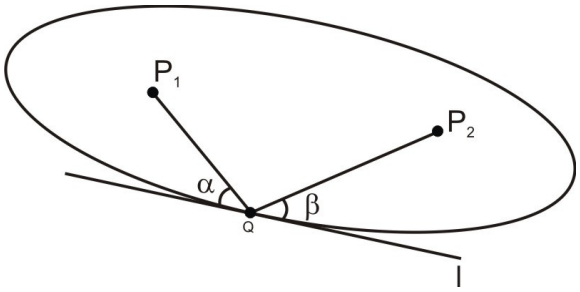
A nice way to find this is to reflect P_2 across L, obtaining the point P_2' in the diagram below:



Since l lies between P_1 and P_2', it's much easier to find the shortest distance between P_1 and P_2' that passes through l. It's simply the shortest path between P_1 and P_2': a straight line! But for every path between P_1 and P_2, we can reflect the part from l to P_2 to get a path equal in length between P_1 and P_2'. So the shortest path between P_1 and P_2 is simply the shortest path between P_1 and P_2' (the straight line), with the part between l and P_2' reflected to get a path between l and P_2. So the shortest path from P_1 to P_2 that intersects with l is P_1Q followed by QP_2.

The part of the above diagram that is going to help us prove the optical property is that since $\angle 1 \cong \angle 2$ (vertical angles) and $\angle 2 \cong \angle 3$ (reflected angles), then $\angle 1 \cong \angle 3$ (transitive property). Thus the shortest path from P_1 to P_2 that intersects with l consists of two segments that meet the line at equal angles.

Now to prove the optical property of the ellipse, apply the above situation to the ellipse. In the picture below, P_1 and P_2 are foci of the ellipse and Q is the intersection with tangent line l.



The optical property states that $\angle\alpha \cong \angle\beta$. This is true because all points other than Q on line l lie outside the ellipse. Points outside the ellipse have a combined distance to the two foci that is greater than points on the ellipse. So Q is the point on line l with the smallest combined distance to points P_1 to P_2. Thus, like we showed in general above, $\angle\alpha \cong \angle\beta$. So the optical property has been proved.

Review Questions

20. Design the largest possible echo room with the following constraints: You would like to spy on someone who will be 3 m from the tip of the ellipse. The room cannot be more than 100 m wide in any direction. How far from the person you're spying on will you be standing.

Review Answers

20. The echo room has a major axis of 100 m and a minor axis of 34.12 m. Situating the room in the coordinate plane, the room can be represented by the equation: $\frac{x^2}{2500} + \frac{y^2}{291} = 1$. You will be 94 m from the person you are spying on.

Sundials:

Figure 1.2

Conic sections help us solve the problem of making a sundial. Depending on the season, the sun shines at a different angle. However, due to the elliptical nature of the Earth's orbit about the sun, a shadow-casting stick can be placed in such a way that the shadow always tells the correct time of day, no matter what the time of year, as long as the stick is lined up with the earth's pole of rotation.

Review Questions

21. No matter what the orientation of a stick, if you trace out the path that the shadow of the tip makes on a flat surface, you will find it is an ellipse. Describe why this is true.(HINT: for simplicity, you can assume that you are making the measurements throughout the course of one day and that with the exception of the earth rotating about a pole, the sun and the earth are fixed with respect to one another.)
22. It was mentioned earlier in the chapter that when a round glass of water is tilted, the surface of the water is an ellipse. Or, in other words, this statement is claiming that the cross section of a cylinder is an ellipse. Prove that this is true. (Hint: to prove that this is an ellipse all you need to do is show that for any cross section of a cylinder there exists a cone that has the same cross section.)
23. From the exercise above, it appears that there is some overlap between "conic sections" and "cylindrical sections". Are any of the classes of conic sections we found in the last section not cylindrical sections? Are there any cylindrical sections that are not conic sections?

Review Answers

21. Answers may vary.
22. Answers may vary.
23. Answers may vary.

Vocabulary

Ellipse A conic section that can be equivalently defined as: 1) any finite conic section, 2) a circle which has been dilated (or "stretched") in one direction, 3) the set of points in which the sum of distances to two special points called the foci is constant.

Major axis The segment spanning an ellipse in the longest direction.

Minor axis The segment spanning an ellipse in the shortest direction.

Focus One of two points that defines an ellipse in the above definition.

Eccentricity A measure of how "stretched out" an ellipse is. Formally, it is the distance between the two foci divided by the length of the major axis. The eccentricity ranges from 0 (a circle) to points close to 1, which are very elongated ellipses.

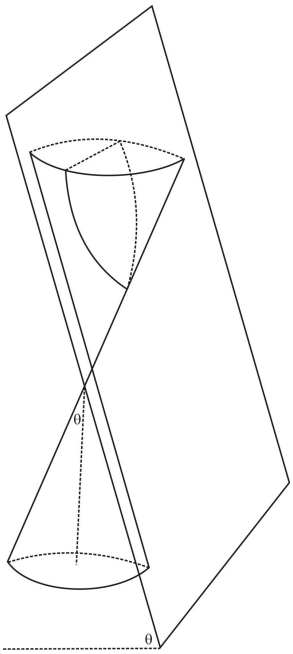
1.3 Parabolas

Learning Objectives

- Understand what results when a cone is intersected by a plane parallel to one side of the cone.
- Understand the focal-directrix property for parabolas.
- Recognize and work with equations for parabolas.
- Understand that all parabolas are self-similar.
- Understand the equivalence of different definitions of parabolas.

Introduction

We've examined ellipses and circles, the two cases when a plane intersects only one side of the cone and creates a finite cross-section. Is it possible for a plane to intersect only one side of the cone, but create an infinite cross-section? It turns out that this is possible if and only if the plane is parallel to one of the lines making up the cone. Or, in other words, the angle θ between the plane and the horizon, is equal to the angle formed by a side of the cone and the horizontal plane.



In the image above, if you till the plane a little bit to the left it will cut off a finite ellipse (possibly a very large one if you only tilt it a little.) Tilt the plane to the right and it will intersect both sides of the cone, making a two-part conic section called a hyperbola, which will be discussed in the next section.

When the plane is parallel to from the side of the cone, the infinite shape that results from the intersection of the plane and the cone is called a **parabola**. Like the ellipse, it has a number of interesting geometric properties.

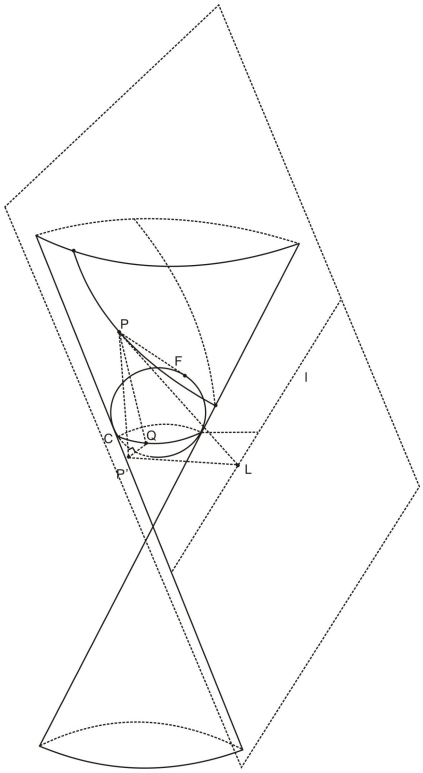
Focus-Directrix Property

Like the ellipse, the parabola has a focal property. And, also like the ellipse, a construction similar to Dandelin's with the spheres can show us what it is. Dandelin himself didn't prove the focal property for parabolas that we are about to discuss, but Pierce Morton used a sphere construction similar to Dandelin's to prove the focal property of parabolas in 1829. We'll look at Morton's argument here.

In contrast with the argument we made for the ellipse, for a parabola we can only fit one tangent sphere inside the cone. That is, only one sphere can be tangent to both the cone and the cutting plane. In the diagram below, the sphere fits underneath the cutting plane, but there is no room for a sphere to lie on top of the cutting plane and still be tangent to the cone.

As with the ellipse, the point where the sphere intersects the plane is called a **focus**. But because there is only one sphere in this construction, and this is related to the fact that a parabola has only one focus. The other geometric object of interest is called the **directrix**. This is the line that results from the intersection between the cutting plane and the plane that contains the circle of contact between the sphere and the cone. In the diagram below, the directrix is labeled l and is found by intersecting the plane defined by

circle C and the cutting plane (the planes are shown in dashed lines for clarity). Finally, we will call the angle between the planes θ.



In the above diagram, we have labeled the point where the sphere contacts the cutting plane with F, and we'll call that point the focus of the parabola. Suppose P is an arbitrarily chosen point on the parabola. Then, let Q be the point on circle c such that $\overline{PQ}$ is tangent to the sphere. In other words Q is chosen so that $\overline{PQ}$ lies on the cone itself. Let L be the point on the directrix l such that $\overline{PL}$ is perpendicular to l. Then $PF = PQ$ since both segments are tangents to the sphere from the same point P. We can also show that $PQ = PL$. This follows from the fact that the cutting plane is parallel to one side of the cone. Consider the point P' that is the projection of P onto the plane containing circle C. Then $\square\ PP'Q$ and $\square\ PP'L$ are both right angles by the definition of a projection. $\square\ PQP'$ and $\square\ PLP'$ are both equal to the angle $90 - \theta$, where θ is the angle defined above, because the cutting plane and the cone both have an angle of θ with the horizon. Since they also share a side, triangles $\triangle PQP'$ and $\triangle PLP'$ are congruent by AAS. So the corresponding sides $\overline{PQ}$ and $\overline{PL}$ are congruent. By the transitive property we have $PF = PL$, so the distance between the point P on the parabola to the focus is the same as the distance between P and the directrix l. We have just proven the focus-directrix property of parabolas.

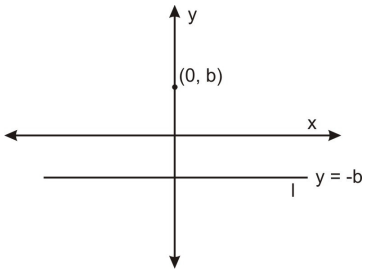
Equations and Graphs of Parabolas

The equation of a parabola is simpler than that of the ellipse. We will arrive at the equation for a parabola in two ways.

Method 1: Using the Distance formula

The first method arises directly from the focus-directrix property discussed in the previous section. Suppose we have a line and a point not on that line in a plane, and we want to find the equation of the set of points in the plane that is equidistant to these two objects. Without losing any generality, we can orient the line

horizontally and the point on the $y-$axis, with the origin halfway between them. Since the parabola is the set of points equidistant from the line and the point, The parabola passes through the origin, (0,0). The picture below shows this configuration. The point directly between the directrix and the focus (the origin in this case) is called the **vertex** of the parabola. Suppose the focus is located at $(0, b)$. Then the directrix must be $y = -b$.



Thus, the parabola is the set of points (x, y) equidistant from the line $y = -b$ and the focus point $(0, b)$. The distance to the line is the vertical segment from (x, y) down to $(0, -b)$, which has length $y - (-b) = y + b$. The distance from (x, y) to the focus $(0, b)$ is distance $= \sqrt{(x-0)^2 + (y-b)^2}$ by the distance formula. So the equation of the parabola is the set of points where these two distances equal.

$$y + b = \sqrt{(x-0)^2 + (y-b)^2}$$

Since distances are always positive, we can square both sides without losing any information, obtaining the following.

$$\begin{aligned} y^2 + 2by + b^2 &= x^2 + y^2 - 2by + b^2 \\ 2by &= x^2 - 2by \\ 4by &= x^2 \\ y &= \frac{1}{4b}x^2 \end{aligned}$$

But b was chosen arbitrarily and could have been any positive number. And for any positive number, a, there always exists a number b such that $a = \frac{1}{4b}$, so we can rewrite this equation as:

$$y = ax^2$$

where a is any constant.

This is the general form of a parabola with a horizontal directrix, with a focus lying above it, and with a vertex at the origin. If a is negative, the parabola is reflected about the $x-$axis, resulting in a parabola with a horizontal directrix, with a focus lying below it, and with a vertex at the origin. The equation can be shifted horizontally or vertically by moving the vertex, resulting in the general form of a parabola with a horizontal directrix and passing through a vertex of (h, k):

$$y - k = a(x - h)^2$$

Switching x and y, the equation for a parabola with a vertical directrix and with a vertex at (h, k) is:

$$x - h = a(y - k)^2$$

Example 1

Sketch a graph of the parabola $y = 3x^2 + 12x + 17$.

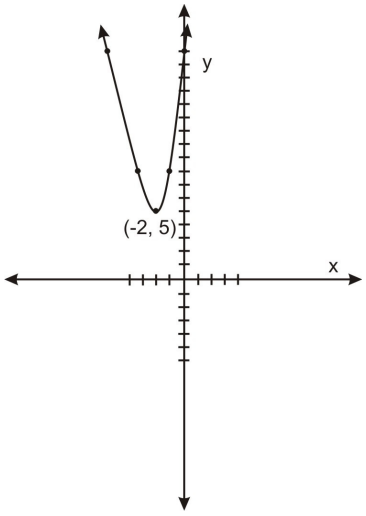
Solution: First, we need to factor out the coefficient of the x^2 term and complete the square:

$$y = 3(x^2 + 4x) + 17$$
$$y = 3(x^2 + 4x + 4) + 17 - 12$$
$$y = 3(x + 2)^2 + 5$$

Now we write it in the form $x - h = a(y - k)^2$:

$$y - 5 = 3(x + 2)^2$$

So the vertex is at (-2,5) and plotting a few x–values on either side of $x = -2$, we can draw the following sketch of the parabola:

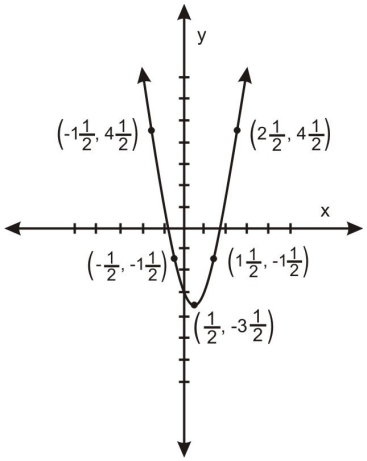


Review Questions

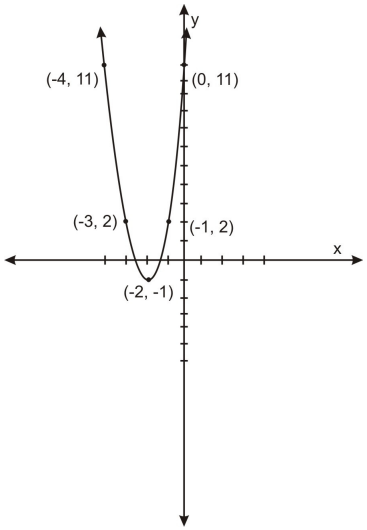
1. Sketch a graph of the following parabola: $y = 2x^2 - 2x - 3$
2. Sketch a graph of the following parabola: $3x^2 + 12x + 11 - y = 0$
3. Sketch a graph of the following parabola: $0 = x^2 - y + 3x + 5$
4. Identify which of the following equations are parabolas: $y - 5x + x^2 = 3$, $x^2 - 3y^2 + 3y - 2x + 15 = 0$, $x - 6y^2 + 20x - 100 = 0$
5. Draw a sketch of the following parabola. Also identify its directrix and focus. $3x^2 + 6x - y = 0$
6. Find the equation for a parabola with directrix $y = -2$ and focus (3,8).

Review Answers

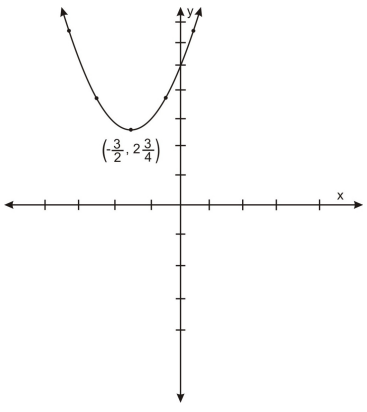
1.



2.

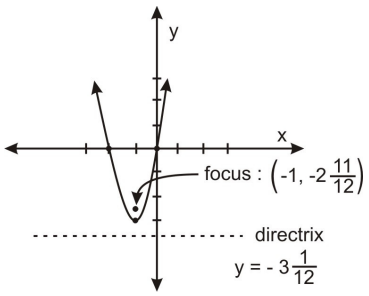


3.



4. $y - 5x + x^2 = 3$ and $x - 6y^2 + 20x - 100 = 0$

5.

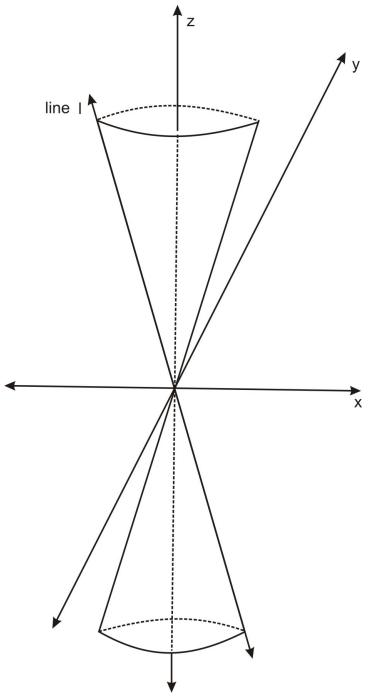


6. $y - 3 = \frac{1}{20}(x-3)^2$

Method 2: Using Three Dimensional Analytical Geometry

Alternatively, we could have arrived at the above equation for a parabola without the argument of Pierce Morton or the use of Dandelin's spheres. The following argument is closer to Apollonius' approach to parabolas in Ancient Greece, though the closed equation forms and the coordinate grid that I will use are modern conventions.

Consider a cone oriented in space as pictured below:



If the cone opens at an angle such that at any point its radius to height ratio is a, then the cone could be defined as the set of points such that the distance from the $z-$axis is a times the $z-$coordinate. Or, in other words, the set of points (x, y, z) satisfying:

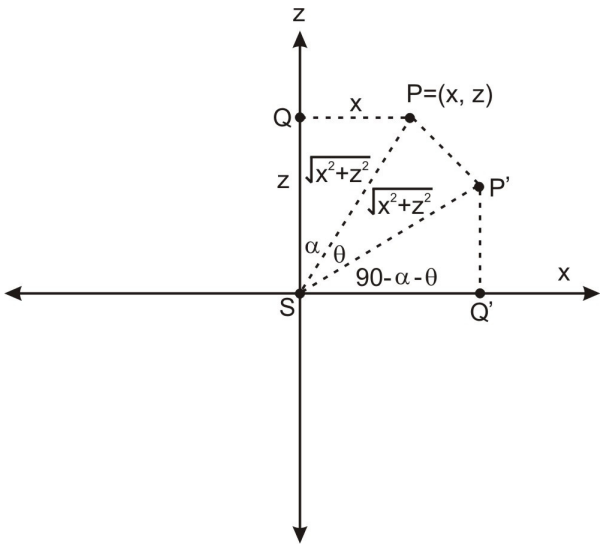
$$\sqrt{(x-0)^2 + (y-0)^2} = az$$

Or:

$$x^2 + y^2 = a^2 z^2$$

This equation works for negative values of x, y, and z, giving the general equation for a two-sided cone.

To consider the intersection of this cone with a plane that is parallel to the line marked l marked in the diagram above, it is most convenient to rotate the entire cone about the y axis until the left side of the cone is vertical, then intersect it with a vertical plane perpendicular to the $x-$axis. Such a rotation leaves the $y-$variable unchanged. To see what it does to the x and z variables, let's see what happens to the point (x, z) on the $xz-$plane when it is rotated by an angle of θ.



In the above diagram, $P(x, z)$ is rotated by an angle of θ to the point P'. We have marked the side lengths $QP = x$ and $SQ = z$. By the Pythagorean Theorem, $SP = \sqrt{x^2 + z^2}$. We also have $SP' = \sqrt{x^2 + z^2}$, since rotation leaves the distance from the origin unchanged. To find the $x-$coordinates of our rotated point P', we can use the fact that $\cos(90 - \alpha - \theta) = \frac{SQ'}{\sqrt{x^2+z^2}}$. But by properties of cosine we have:

$$\cos(90 - \alpha - \theta) = \sin(\alpha + \theta),$$

and substituting with the sine addition formula gives us:

$$\frac{SQ'}{\sqrt{x^2 + z^2}} = \sin(\alpha)\cos(\theta) + \cos(\alpha)\sin(\theta),$$

which we can use our diagram to change to:

$$\frac{SQ'}{\sqrt{x^2 + z^2}} = \frac{x}{\sqrt{x^2 + z^2}}\cos(\theta) + \frac{z}{\sqrt{x^2 + z^2}}\sin(\theta)$$

which simplifies to:

$$SQ' = x\cos(\theta) + z\sin(\theta)$$

To find the $x-$coordinates of our rotated point P', we can use the fact that $\sin(90 - \alpha - \theta) = \frac{P'Q'}{\sqrt{x^2+z^2}}$. But by properties of sine we have:

$$\sin(90 - \alpha - \theta) = \cos(\alpha + \theta)$$

and substituting with the cosine addition formula gives us:

$$\frac{P'Q'}{\sqrt{x^2 + z^2}} = \cos(\alpha)\cos(\theta) - \sin(\alpha)\sin(\theta),$$

which we can use our diagram to change to:

$$\frac{P'Q'}{\sqrt{x^2+z^2}} = \frac{z}{\sqrt{x^2+z^2}}\cos(\theta) - \frac{x}{\sqrt{x^2+z^2}}\sin(\theta)$$

which simplifies to:

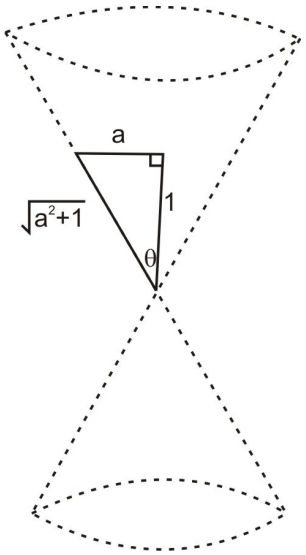
$$P'Q' = z\cos(\theta) - x\sin(\theta)$$

Looking back at the picture, this means that the coordinates of P' are $(x\cos(\theta)+z\sin(\theta),\ z\cos(\theta)-x\sin(\theta))$. In other words, in rotating from P to P', the $x-$coordinate changes to $x\cos(\theta)+z\sin(\theta)$ and the $z-$ coordinate changes to $z\cos(\theta) - x\sin(\theta)$.

If this rotation happens to every point on the cone, we can substitute $x\cos(\theta)+z\sin(\theta)$ for x and $z\cos(\theta) - x\sin(\theta)$ for z into our equation of the cone, resulting in a new equation for the cone after rotating by θ.

$$\begin{aligned}
(x\cos(\theta)+z\sin(\theta))^2 + y^2 &= a^2(z\cos(\theta) - x\sin(\theta))^2\\
x^2\cos^2(\theta) + 2xz\cos(\theta)\sin(\theta) + z^2\sin^2(\theta) + y^2 &= a^2(x^2\sin^2(\theta) - 2xz\cos(\theta)\sin(\theta) + z^2\cos^2(\theta))\\
x^2\cos^2(\theta) + 2xz\cos(\theta)\sin(\theta) + z^2\sin^2(\theta) + y^2 &= a^2x^2\sin^2(\theta) - 2a^2xz\cos(\theta)\sin(\theta) + a^2z^2\cos^2(\theta))\\
x^2\cos^2(\theta) + 2xz\cos(\theta)\sin(\theta) + z^2\sin^2(\theta) + y^2 &= a^2x^2\sin^2(\theta) - 2a^2xz\cos(\theta)\sin(\theta) + a^2z^2\cos^2(\theta))
\end{aligned}$$

Now in the case of the tilted cone, we want to tilt the cone such that the left side becomes vertical. Since the factor a determines how tilted the cone is, we can see from the triangle below that $\sin(\theta) = \frac{a}{\sqrt{1+a^2}}$ and $\cos(\theta) = \frac{1}{\sqrt{1+a^2}}$.

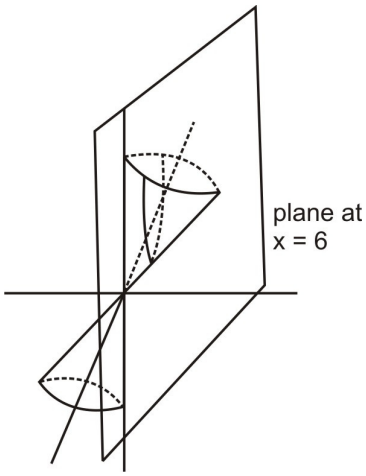


So the equation becomes:

$$\begin{aligned}
x^2\frac{1}{1+a^2} + 2xz\frac{a}{1+a^2} + z^2\frac{a^2}{1+a^2} + y^2 &= a^2x^2\frac{a^2}{1+a^2} - 2a^2xz\frac{a}{1+a^2} + a^2z^2\frac{1}{1+a^2}\\
x^2 + 2xza + z^2a^2 + y^2(1+a^2) &= a^4x^2 - 2a^3xz + a^2z^2\\
x^2 + 2xza + y^2(1+a^2) &= a^4x^2 - 2a^3xz
\end{aligned}$$

Now that we have tilted our cone, to take a cross section that is parallel to the left side of the cone, we can simply cut it with a vertical plane. The equation of a vertical plane going through $(b,0,0)$ and

perpendicular to the $x-$axis is $x = b$. Therefore, setting x equal to the constant in the equation above will give us the intersection of the tilted cone and a plane parallel to one side of the cone. b. Here is a picture of the rotation and the cross-section, which lies in an $xz-$plane.



Setting x equal to the constant b, we have:

$$\begin{aligned} b^2 + 2abz + y^2(1 + a^2) &= a^4b^2 - 2a^3bz \\ z(2ab + 2a^3b) &= -y^2(1 + a^2) + a^4b^2 - b^2 \\ z &= \left(\frac{-1 - a^2}{2ab + 2a^3b}\right)y^2 + (a^4b^2 - b^2) \end{aligned}$$

Although this coefficient and constant term seem complicated, a and b can be chosen so that the coefficient of the y^2 term can be equal to any number (you will explore this fact in an exercise). The constant term can be ignored since any parabola can be shifted vertically by any amount.

So the general form of a parabola is:

$$z = Ay^2$$

where A is any constant.

Or, using the more standard $x-$ and $y-$coordinates the form of a parabola is

$$y = ax^2$$

As before, this equation can be adapted to produce the shifted and horizontally oriented forms.

Review Questions

7. Explain why $\cos(90 - \alpha - \theta) = \sin(\alpha + \theta)$ in the above argument.
8. Show that for any A there exist constants a and b such that $A = \frac{-1-a^2}{2ab+2a^3b}$.

Review Answers

7. There are many arguments that work. One route is to use the fact that $\cos(\alpha) = \cos(-\alpha)$ for any α, and then the fact that $\cos(\alpha - 90) = \sin(\alpha)$ for any α.

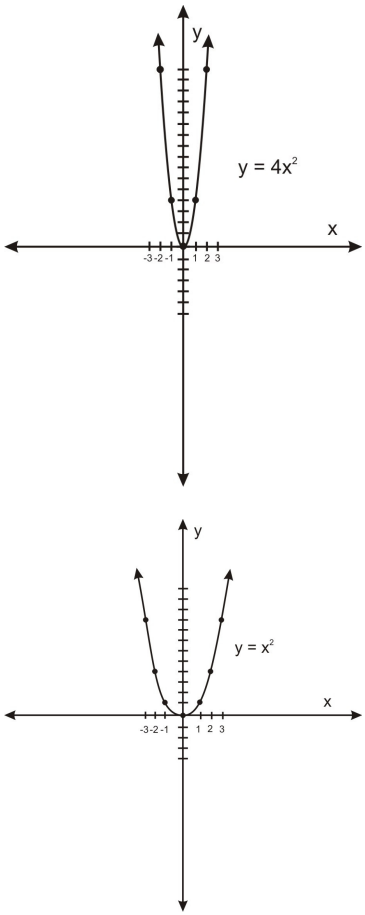
8. Solving for b in terms of A and a, we have:

$$\begin{aligned} A(2ab + 2a^3b) &= -1 - a^2 \\ 2Aab(1 + a^2) &= -(1 + a^2) \\ 2Aab &= -1 \\ 2Aab &= -1 \\ b &= -\frac{1}{2Aa} \end{aligned}$$

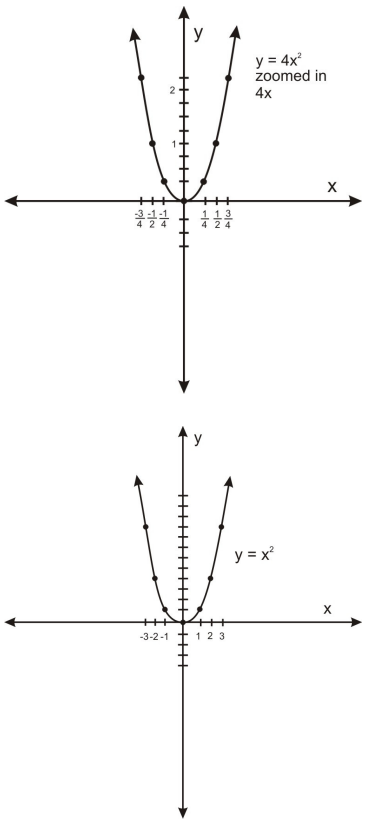
So we can set $b = 2Aa$ and the relationship will hold.

All Parabolas are the Same

There is a subtle point in the above two arguments that reveals a very interesting property of parabolas. This is the fact that *all parabolas have the same shape.* Or, in the language of geometry, any two parabolas are similar to one another. This means that one parabola can be scaled in or out to produce another parabola of exactly the same shape. This may come across as surprising, because parabolas where x^2 has a large coefficient certainly look much "steeper" than parabolas with a small coefficient when examined over the same domain, as shown in the graphs below.

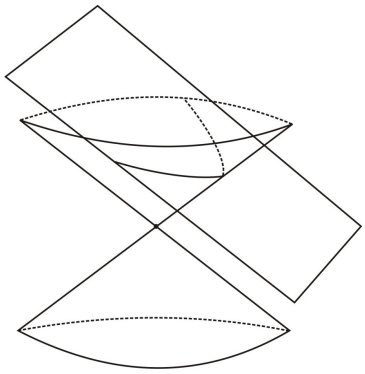


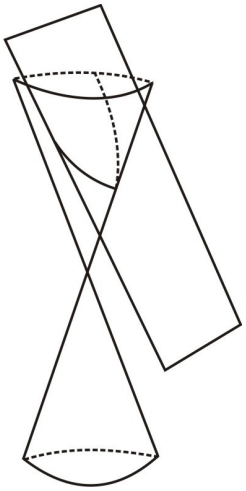
But when one of the parabolas is scaled appropriately, these parabolas are identical:



This fact about parabolas can be seen in the first argument simply from the fact that all parabolas are generated from a line and a point not on that line. This configuration of generating objects, a line and a point, is always the same shape. Any other line and point looks exactly the same—simply zoom in or out until the line and point are the same distance from one another. So the shapes that any two such configurations generate must also be the same shape.

In the second argument, there were two factors that might affect the shape of the parabola. The first is the distance between the cutting plane and the apex of the cone. But cones have the same proportions at any scale, so no matter what this distance, the picture can be reduced or enlarged, affecting this distance but not the shape of the cone or plane. So this parameter does not actually change the shape of the conic section that results. The other factor is the shape of the actual cone. This is its steepness, defined by the angle at the apex, or equivalently by the ratio between the radius and the height at any point. This is a bit trickier. It's not at all obvious that short, squat cones and tall, skinny cones would produce parabolas of the same shape.





According to what we found, any parabola produced by slicing any cone resulted in an equation of this form:

$$y = ax^2$$

We want to show that if we generate two such parabolas, that they actually have the same shape. So suppose we use two cone constructions and come up with these parabolas: $y = a_1x^2$ and $y = a_2x^2$. We want to show there is some scale factor, call it f, that shrinks or enlarges $y = a_1x^2$ into $y = a_2x^2$. To keep a shape the same, the scale factor needs to affect both the $x-$ and $y-$variables. So we need to find an f such that $(fy) = a_1(fx)^2$ is equivalent to $y = a_2x^2$. The first equation can be written $y = (a_1f)x^2$, which is equivalent to the second equation when $a_1f = a_2$, or when $f = \frac{a_2}{a_1}$. Such an f always exists for non-negative numbers a_1 and a_2. So the parabolas are indeed the same shape. If a is less than zero, then the parabola can be reflected vertically to produce a parabola of the same shape and with positive coefficient a.

Review Questions

9. What other classes of shapes have this property that all members of the class are similar to each other?
10. Explain why not all ellipses are similar. While enlarging or shrinking doesn't work to make two ellipses identical, how can you change the view of two ellipses that have different shapes so that they look the same?

Review Answers

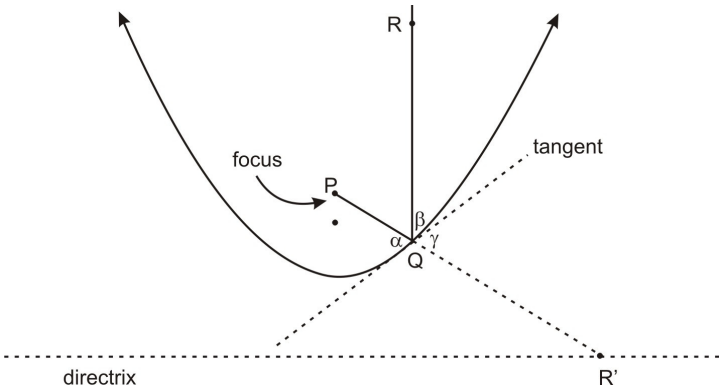
9. A few examples are: circles, squares, finite sections of one-sided cones of the same angle.
10. The eccentricity of ellipses defines the shape, so when the eccentricity is different for two ellipses, the ellipses are not similar to one another. Viewing one of the ellipses at an angle, however, changes the perceived eccentricity of that ellipse, and the angle can be chosen to match the perceived eccentricity to the eccentricity of the other ellipse, producing an image that is similar to the other ellipse.

Applications

"Burning Mirrors"

Diocles (~ 240 – 180 BCE) was a mathematician from Ancient Greece about whom we know very little. However, we know enough from a few scant documents that he thought about an important application of parabolas. It comes from what is sometimes called the "optical property" of parabolas. Remember the optical property of ellipses: lines from one focus "bounce off" the side of the ellipse to hit the other focus.

For parabolas, since parabolas have only one focus, the directrix plays a role. For the parabola, the optical property is that lines perpendicular to the directrix "bounce off" the parabola and converge at the focus. Or, alternatively, lines from the focus "bounce off" off the parabola and continue perpendicular to the directrix. As with the ellipse, "bouncing off" means that the two lines meet the parabola at equal angles to the tangent.



In the above diagram, the optical property states that $\square\ \alpha \cong \square\ \beta$. The proof is similar to the proof of the optical property of ellipses. In the above diagram P is the focus and Q is a point on the parabola. Let R' be the point on the directrix that is obtained by extending PQ. Then PR', the straight line, is clearly the shortest distance between P and R' that passes through the tangent line. Let R be on the line lying directly above Q such that $QR = QR'$. The R can be thought of as R', reflected across the tangent line. Then $\square\ \alpha \cong \square\ \gamma$ (vertical angles) and $\square\ \gamma \cong \square\ \beta$ (reflected angles), and so $\square\ \alpha \cong \square\ \beta$ (transitive property).

The optical property has some interesting applications. Diocles described one potential application in his document "On Burning Mirrors". He envisioned a parabolic-shaped mirror (basically a parabola rotated about its line of symmetry) which would collect light from the sun and focus it on the focal point, creating enough of a concentration of light to start a fire at that point. Some claim that Archimedes attempted to make such a contraption with copper plates to fight the Romans in Syracuse.

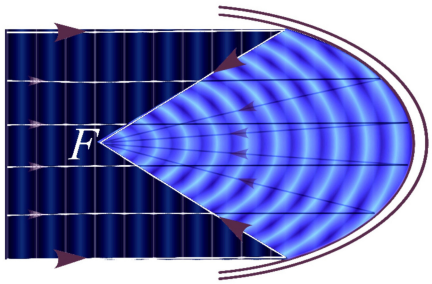


Figure 1.3

Review Questions

11. If nothing was used to deflect light before it entered the mirror, where would the sun have to be in relationship to you and the place you want to start a fire? Why is this a constraint? Design a way to circumvent this constraint.
12. You live in Ancient Greece and are defending your city against Roman invaders. Design a "burning mirror" that you plan to use to destroy an army that is approaching at a distance of three miles.

Review Answers

11. The fire-locale must lie on the segment between you and the sun. This is a problem because to start a ground fire, you would have to wait until evening when the sun is not as bright. A lens or mirror that changes the angle of the suns rays could help you work around this constraint.
12. Answers will vary. The distance from the focus to the vertex should be 3 miles.

Headlights

The optical property is also responsible for parabola-shaped unidirectional lights, such as car headlights. If a bulb is placed at the focus of a parabolic mirror, the light rays reflect off the mirror parallel to each other, making a focused beam of light.

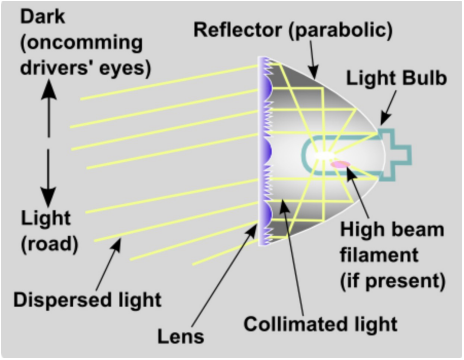


Figure 1.4

Review Questions

13. In the above diagram of a car headlight, the lens directs the beams of light downwards, to keep them out of the eyes of oncoming drivers. But if that was the only purpose of the lens, alternatively the lens could be omitted and the headlight could just be angled down slightly. But there is another purpose to the lens. What is it?

Review Answers

13. The lens also expands the array of light which is why it is called "dispersed light." Without the lens, the headlight would only illuminate a strip the width of the headlight itself, which would not be very useful for driving.

Cassegrain Telescopes

Satellite telescopes take advantage of the optical property of parabolas to collect as much light from a distant star as possible. The dish of the satellite below is parabolic in shape and reflects light to the point in the middle.

Figure 1.5

This image at this point is then focused with a lens into the telescope as shown in the diagram below. (from:http://commons.wikimedia.org/wiki/File:Telescope_cassegrain_principe.png) The

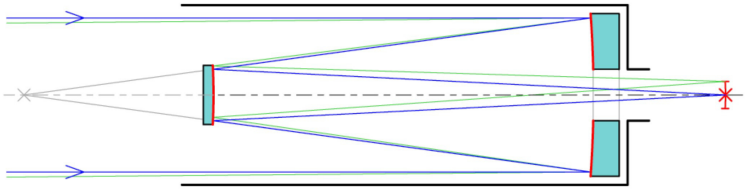


Figure 1.6

Technology

As with ellipses, you can examine the graphs of parabolas on a graphing calculator or computer program. But as with ellipses, there is a much more low-technology tool that can draw a parabola. It is a generalized version of the circle-drawing compass. Below is a model of a conic section producing compass drawn by Leonardo da Vinci. Like a compass for making circles, this tool assists a pencil in swiveling about a vertex. Unlike a compass, the pencil is held loosely in a shaft so that it can slide up and down, and the fixed side is held at a constant angle. As this angle changes, different conic sections result. This essentially turns the pencil shaft into the cone and the paper into the cutting plane.

Review Questions

14. Describe how the above compass should be set up to produce a parabola.

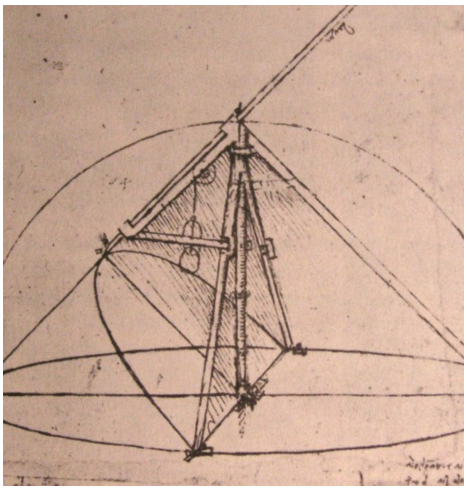


Figure 1.7

Review Answers

14. The pencil should be parallel to the paper at its most extended point.

Vocabulary

Parabola A conic section resulting from intersecting a cone with a plane that is parallel to one of the lines in the cone.

Focus The point to which all points on a parabola are the same distance as they are to a line, called the directrix.

Directrix A line to which all points on a parabola are the same distance as they are to a point called the focus.

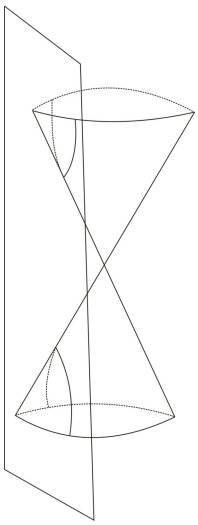
1.4 Hyperbolas

Learning Objectives

- Understand what results when a cone is intersected by a plane that intersects both sides of the cone.
- Understand the focal-directrix property for hyperbolas.
- Recognize and work with equations for hyperbolas.
- Understand how a hyperbola can be described in terms of its asymptotes and foci.
- Understand the equivalence of different definitions of hyperbolas.

Two-part Conic Sections

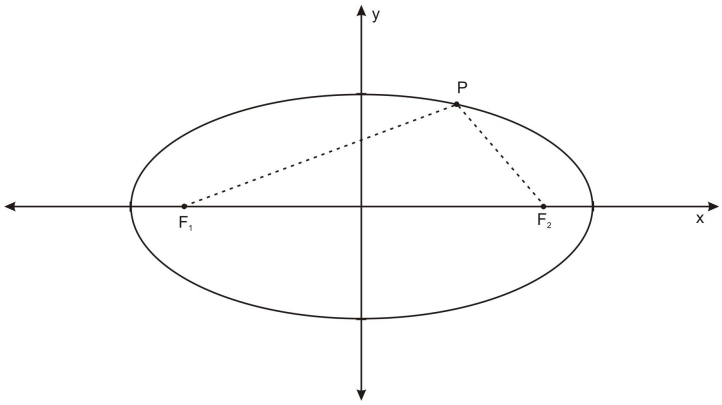
The final way of cutting a cone with a plane appears at first to be much messier. Not only does it result in an infinite shape, but there are two pieces that aren't even connected! When the plane slices through two parts of the cone, the two infinite "U"-shaped parts are together called a **hyperbola**.



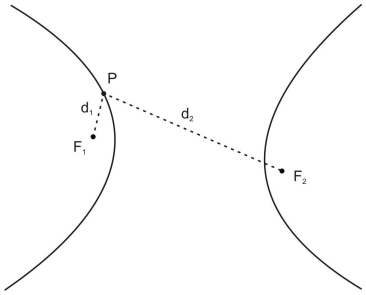
In this section we will see that this sprawling shape actually has some beautiful properties that make it as noble as its cousins.

The Focal Property

Even though this shape seems much harder to conceive of than an ellipse, the hyperbola has a defining focal property that is as simple as the ellipse's. Remember, an ellipse has two foci and the shape can be defined as the set of points in a plane whose distances to these two foci have a fixed sum.



Hyperbolas also have two foci, and they can be defined as the set of points in a plane whose distances to these two points have the same *difference*. So in the picture below, for every point P on the hyperbola, $|d_2 - d_1| = C$ for some constant C.



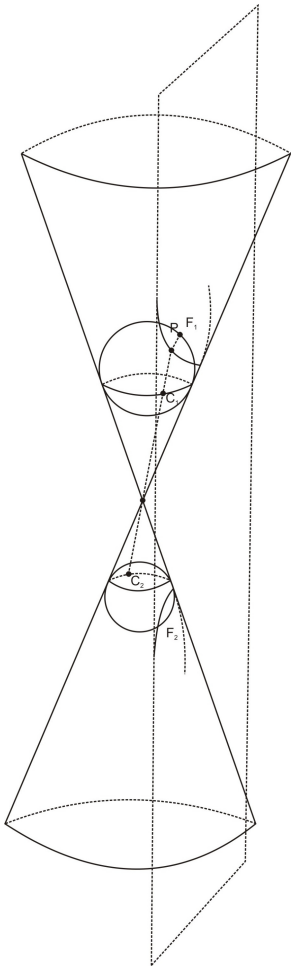
Review Questions

1. The hyperbola is infinite in size. In mathematics this is called **unbounded**, which means no circle, no matter how large, can enclose the shape. Explain why a focal property involving a *difference* results in an unbounded shape, while a focal property involving a sum results in a bounded shape.

Review Answers

1. Answer should include the following concept: In the case of an ellipse, we had two distances summing to a constant. Since the distances are both positive then there is a limit to the size of the numbers. In the case of hyperbolas, two very large positive numbers can have a much smaller difference.

To prove the focal property of hyperbolas, we examine Dandelin's sphere construction. Unlike the construction for ellipses, which used two spheres on one side of the cone, and the sphere construction for parabolas, which used one sphere on one side of the cone, this construction uses two spheres, one on each side of the cone. As with the ellipse construction, each sphere touches the plane at one of the foci of the hyperbola. And as with the argument for the elliptical focal property, the argument uses the fact that tangents from a common point to a sphere are equal.



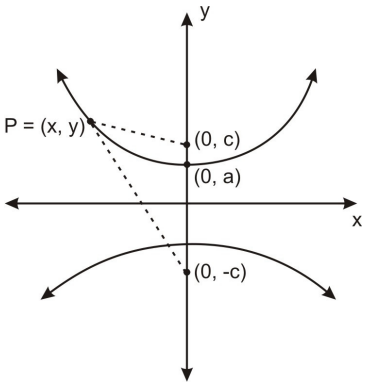
In the above diagram, suppose P is an arbitrary point on the hyperbola. We would like to examine the difference $PF_2 - PF_1$. Let C_1 be the point on the upper sphere that lies on the line between P and the vertex of the cone. Let C_2 be the point on the upper sphere when this line is extended (so P, C_1, and

C_2 are all on the same line and $PC_1 + C_1C_2 = PC_2$ and the cone By the common tangent property, $PF_1 = PC_1$ and $PF_2 = PC_2$ for some points C_1 and C_2 on the circles where the spheres meet the cone. So $PF_2 - PF_1 = PC_2 - PC_1 = (PC_1 + C_1C_2) - PC_1 = C_1C_2$. But C_1C_2 is the distance along the cone between the two circles of tangency and is constant regardless of the choice of C_1 and C_2. So the difference $PF_2 - PF_1$ is constant.

Equations for Hyperbolas

To derive the equation for a hyperbola, we can't do exactly what we did for ellipses. Remember, when we first derived the equations for ellipses, we were defining them as "stretched out circles." Hyperbolas aren't as obviously a simple dilation of a shape as basic as a circle. Instead, we'll use the focal property to derive their equation. This was what we said could be done for ellipses when I said "simply read the proof of the focal property backwards." To derive the equation form of hyperbolas from the focal property, I'll actually show you the steps in the correct order.

Suppose we have a curve (actually a pair of curves), satisfying the focal property for hyperbolas. Let's orient the hyperbola so that the two foci are on the $y-$axis, and equidistant from the origin. In the diagram below the foci are labeled with the points $(0, c)$, and $(0, -c)$.



The focal property states that the difference in distances between an arbitrary point $P = (x, y)$ on the hyperbola and $(0, c)$, and $(0, -c)$ is a constant. In particular, we know that this constant can be computed from any point on the hyperbola, for instance the point on the $y-$axis marked $(0, a)$. The distance between $(0, a)$ and $(0, -c)$ is $a + c$ and the distance between $(0, a)$ and $(0, c)$ is $c - a$. The difference between these two quantities is $(c + a) - (c - a) = 2a$. So $2a$, the distance between the two $y-$intercepts of the hyperbola, is the constant in the focal property for the hyperbola. Using this distance formula and the focal property, we have:

$$\sqrt{(x-0)^2+(y-(-c))^2}-\sqrt{(x-0)^2+(y-c)^2}=2a$$
$$x^2+(y+c)^2-2\sqrt{x^2+(y+c)^2}\sqrt{x^2+(y-c)^2}+x^2+(y-c)^2=4a^2$$
$$2x^2+(y+c)^2+(y-c)^2-4a^2=2\sqrt{x^2+(y+c)^2}\sqrt{x^2+(y-c)^2}$$
$$(2x^2+2y^2+2c^2-4a^2)^2=4(x^2+(y+c)^2)(x^2+(y-c)^2)$$
$$4x^4+8x^2y^2+8c^2x^2+8c^2y^2+4y^4-16a^2x^2+4c^4-16a^2c^2-16a^2y^2+16a^4=4x^4+4x^2(y+c)^2+4x^2(y-c)^2+4(y+c)$$
$$4x^4+8x^2y^2+8c^2x^2+8c^2y^2+4y^4-16a^2x^2+4c^4-16a^2c^2-16a^2y^2+16a^4=4x^4+4x^2(y+c)^2+4x^2(y-c)^2+4(y^2-c$$
$$4x^4+8x^2y^2+8c^2x^2+8c^2y^2+4y^4-16a^2x^2+4c^4-16a^2c^2-16a^2y^2+16a^4=4x^4+8x^2y^2+8c^2x^2+4y^4-8c^2y^2+4c^4$$
$$16c^2y^2-16a^2x^2-16a^2c^2-16a^2y^2+16a^4=0$$
$$c^2y^2-a^2x^2-a^2c^2-a^2y^2+a^4=0$$
$$c^2y^2-a^2y^2-a^2x^2=a^2c^2-a^4$$
$$y^2(c^2-a^2)-x^2(a^2)=a^2(c^2-a^2)$$
$$\frac{y^2}{a^2}-\frac{x^2}{c^2-a^2}=1$$

Like the collapse we witnessed when investigating the focal property for ellipses, this equation became pretty simple. We can even make it simpler. For any positive number b and a, there exists a c such that $b^2=c^2-a^2$. Thus the general form for a hyperbola that open upwards and downwards and whose foci lie on the $y-$axis is:

$$\frac{y^2}{a^2}-\frac{x^2}{b^2}=1$$

Switching x and y we have hyperbolas that open rightwards and leftwards and whose foci lie on the $x-$axis.

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$

These equations have so far been hyperbolas that are centered about the origin. For a hyperbola that is centered around the point (h,k) we have the shifted equations:

$$\frac{(y-k)^2}{a^2}-\frac{(x-h)^2}{b^2}=1$$

for a hyperbola opening up and down, and

$$\frac{(x-h)^2}{a^2}-\frac{(y-k)^2}{b^2}=1$$

For a hyperbola opening left and right.

Example 1

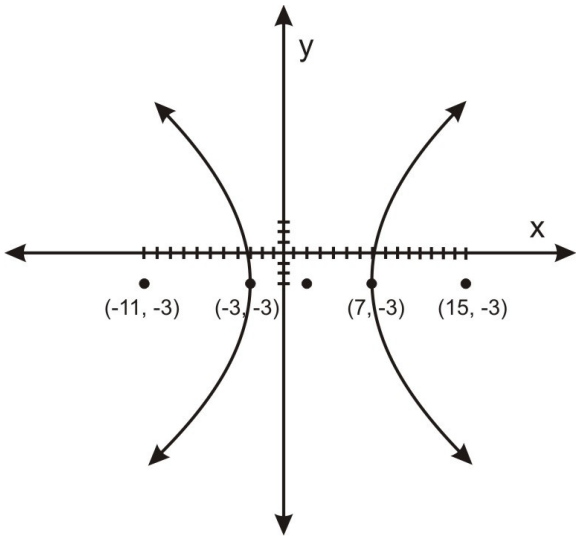
Show that the following equation is a hyperbola. Graph it, and show its foci.

$$144x^2-576x-25y^2-150y-3249=0$$

Solution: The positive leading coefficient for the x^2 term and the negative leading coefficient for the y^2 term indicate that this is a hyperbola that is horizontally oriented. Grouping and completing the square, we have:

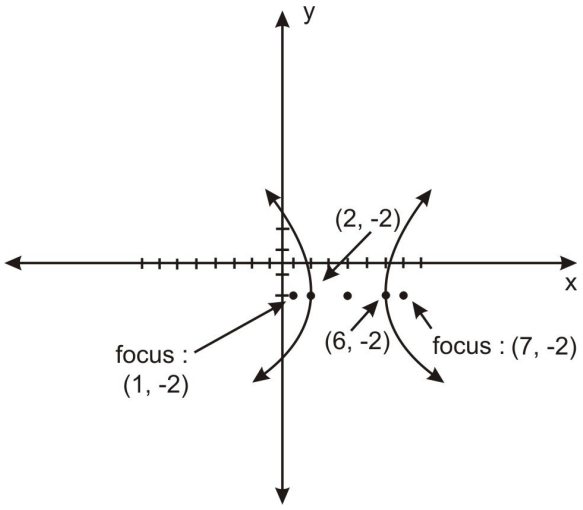
$$144(x^2 - 4x) - 25(y^2 + 6y) = 3249$$
$$144(x^2 - 4x + 4) - 25(y^2 + 6y + 9) = 3249 + 576 - 225$$
$$144(x - 2)^2 - 25(y + 3)^2 = 3600$$
$$\frac{(x-2)^2}{5^2} - \frac{(y+3)^2}{12^2} = 1$$

So our hyperbola is centered at (2,-3). Its vertices are 5 units to the right and left of (2,-3), or at the points (7,-3) and (-3,-3). It opens to the right and left from these vertices. It's foci are c units to the left and right of (2,-3), where $c = \sqrt{a^2 + b^2} = \sqrt{5^2 + 12^2} = 13$. So it's foci are at (15,-3) and (-11,-3). Plotting a few points near (7,-3) and (-3,-3), the graph looks like:

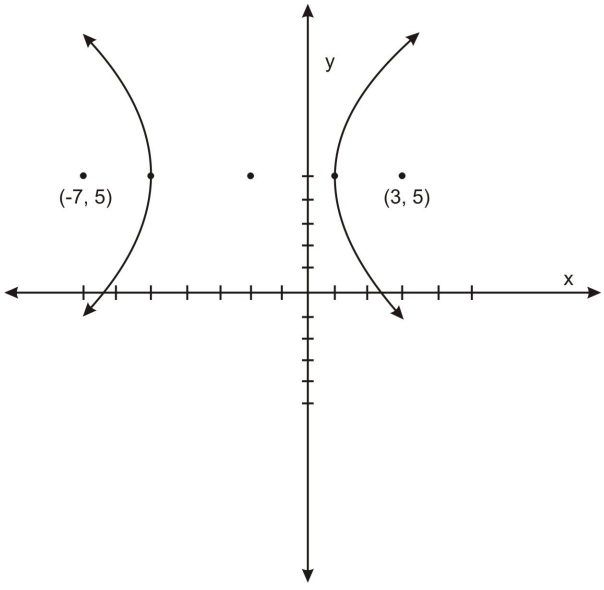


Review Questions

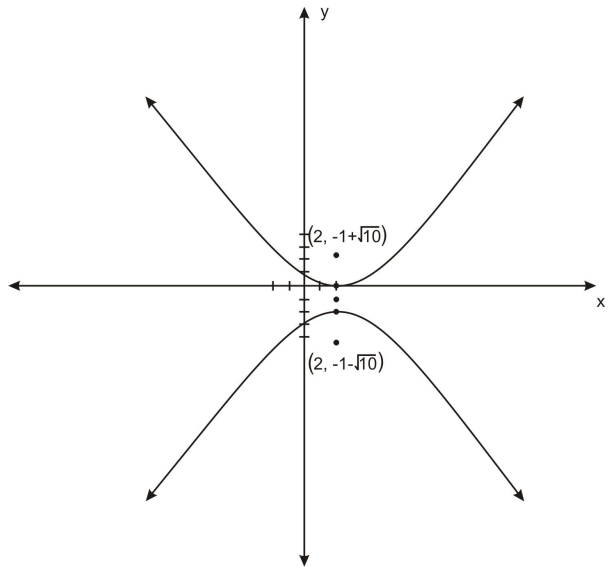
2. Explain why for any positive number b and a, there exists a c such that $b^2 = c^2 - a^2$.
3. Graph the following hyperbola and mark its foci: $16x^2 + 64x - 9y^2 + 90y - 305 = 0$
4. Graph the following hyperbola and mark its foci: $9y^2 + 18y - x^2 + 4x - 4 = 0$
5. Graph the following hyperbola and mark its foci: $25x^2 + 150x - 4y^2 + 24y + 89 = 0$
6. Find the equation for the following hyperbola:



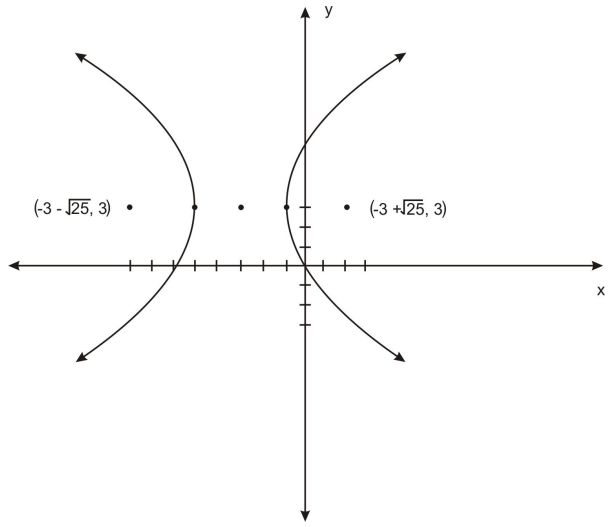
Review Answers

2. Let $c = \sqrt{a^2 + b^2}$. Since $a^2 + b^2$ is always positive for positive a and b, this number is always defined. Geometrically, let c be the hypotenuse of a right triangle with side lengths a and b.
3. 

4.



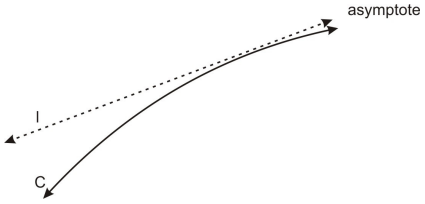
5.



6. $\frac{(x-4)^2}{4} - \frac{(y+2)^2}{45} = 1$

Asymptotes

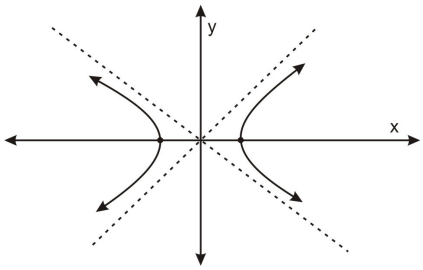
In addition to their focal property, hyperbolas also have another interesting geometric property. Unlike a parabola, a hyperbola becomes infinitesimally close to a certain line as the $x-$ or $y-$coordinates approach infinity. Such a line is called an **asymptote**.



Before we try to prove this property of hyperbolas, we have to figure out what we mean by "infinitesimally close." Here we mean two things: 1) The further you go along the curve, the closer you get to the asymptote, and 2) If you name a distance, no matter how small, eventually the curve will be that close to the asymptote. Or, using the language of limits, as we go further from the vertex of the hyperbola the limit of the distance

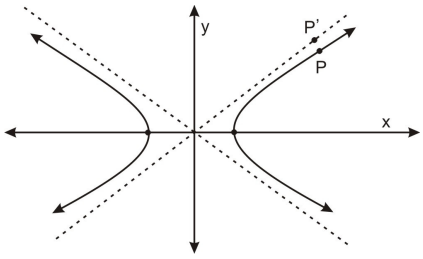
between the hyperbola and the asymptote is 0.

So now we need to prove such a line (or lines) exist for the hyperbola. It turns out there are two of them, and that they cross at the point at which the hyperbola is centered:



It also turns out that for a hyperbola of the form $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, the asymptotes are the lines $y = \frac{b}{a}x$ and $y = -\frac{b}{a}x$. (for a hyperbola of the form $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$ the asymptotes are the lines $y = \frac{a}{b}x$ and $y = -\frac{a}{b}x$. For a shifted hyperbola, the asymptotes shift accordingly.)

To prove that the hyperbola actually gets infinitesimally close to these lines as it goes to infinity, let's focus on the point P on the upper right leg of the hyperbola in the picture below.



Now consider the point P', the point on the asymptote lying directly above P. If we can show that P and P' become infinitesimally close, then that's good enough (this subtlety will be explored in the next exercise.) Well, the point P is on the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, so if it's $x-$value is x, it's $y-$ value can be found by solving for y in the equation. So

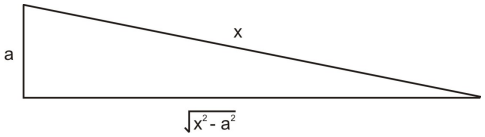
$$\begin{aligned}\frac{x^2}{a^2} - 1 &= \frac{y^2}{b^2} \\ y^2 &= \frac{b^2}{a^2}x^2 - b^2 \\ y &= \sqrt{\frac{b^2}{a^2}x^2 - b^2}\end{aligned}$$

This expression simplifies further, but we'll leave it for now and work on P'. Since P' lies directly above P, x is also the $x-$ coordinate of P'. Similarly as with P, we can find the $y-$coordinate from the equation of the asymptote. Since $y = \frac{b}{a}x$ is the equation of the asymptote, we have $\frac{b}{a}x$ as the $y-$coordinate of P'. So $P = \left(x, \sqrt{\frac{b^2}{a^2}x^2 - b^2}\right)$, and $P' = \left(x, \frac{b}{a}x\right)$. Since these two points are vertically aligned, the distance between them points is simply the difference between the $y-$coordinate:

Distance between P and $P' = \frac{b}{a}x - \sqrt{\frac{b^2}{a^2}x^2 - b^2}$

$$= \frac{b}{a}x - \sqrt{\frac{b^2}{a^2}x^2 - \frac{b^2}{a^2}a^2}$$
$$= \frac{b}{a}x - \sqrt{\frac{b^2}{a^2}(x^2 - a^2)}$$
$$= \frac{b}{a}x - \frac{b}{a}\sqrt{(x^2 - a^2)}$$
$$= \frac{b}{a}\left(x - \sqrt{(x^2 - a^2)}\right)$$

So the distance is a constant multiplied by a quantity that depends on x. What happens to this quantity when x approaches infinity? This kind of question involves working with limits in a way you will study soon. For now, we can use a nice geometric picture to see what happens. In the picture below, we see that a, x, and $\sqrt{(x^2 - a^2)}$ can be thought of as side-lengths of a right triangle, where x is the hypotenuse. This can be verified with the Pythagorean Theorem.



In this picture, what happens to difference in lengths between the side of length x and the side of length $\sqrt{(x^2 - a^2)}$ when x approaches infinity and a remains fixed? The two side lengths become closer in length, so their difference approaches zero.

Going back to the previous diagram where we defined P and P', this means that P and P' become infinitesimally close as x approaches infinity.

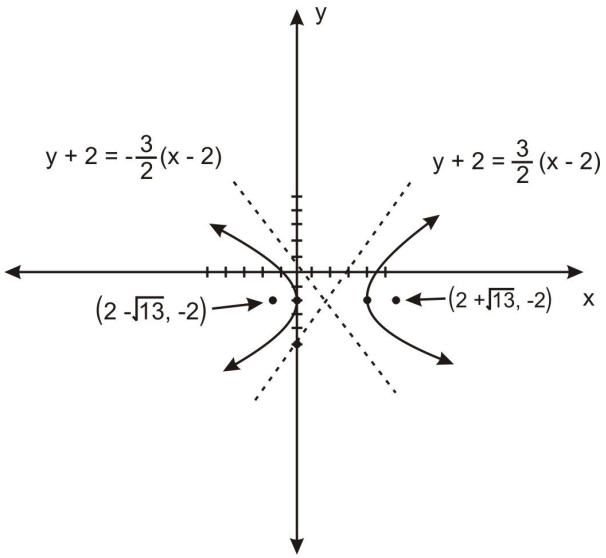
This same argument can be repeated to show that the other three legs of the hyperbola approach their respective asymptotes, and to show that the same holds in vertically-oriented and/or shifted parabolas.

Example: Graph the following hyperbola, drawing its foci and asymptotes and using them to create a better drawing: $9x^2 - 36x - 4y^2 - 16y - 16 = 0$.

Solution: First, we put the hyperbola into the standard form:

$$9(x^2 - 4x) - 4(y^2 + 4y) = 16$$
$$9(x^2 - 4x + 4) - 4(y^2 + 4y + 4) = 36$$
$$\frac{(x-2)^2}{4} - \frac{(y+2)^2}{9} = 1$$

So $a = 2$, $b = 3$ and $c = \sqrt{4 + 9} = \sqrt{13}$. The hyperbola is horizontally oriented, centered at the point (2,-2), with foci at $\left(2 + \sqrt{13}, -2\right)$ and $\left(2 - \sqrt{13}, -2\right)$. After taking shifting into consideration, the asymptotes are the lines: $y + 2 = \frac{3}{2}(x - 2)$ and $y + 2 = -\frac{3}{2}(x - 2)$. So graphing the vertices and a few points on either side, we see the hyperbola looks something like this:



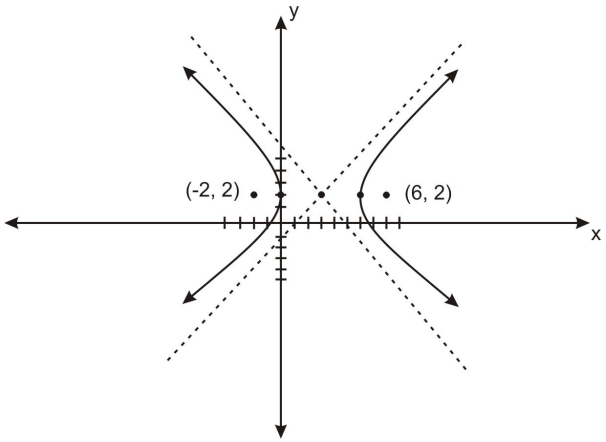
Review Questions

7. In the diagram above, which was used in the proof that hyperbolas have asymptotes, why is it true that if P and P' become infinitesimally close, that this implies the hyperbola and the line become infinitesimally close. (HINT: what exactly do we mean by the "closeness" of a point on a hyperbola and a line?)
8. Graph the following hyperbola, drawing its foci and asymptotes and using them to create a better drawing: $16x^2 - 96x - 9y^2 - 36y - 84 = 0$.
9. Graph the following hyperbola, drawing its foci and asymptotes and using them to create a better drawing: $y^2 - 14y - 25x^2 - 200x - 376 = 0$.
10. Find the equation for a hyperbola with asymptotes of slopes $\frac{5}{12}$ and $-\frac{5}{12}$, and foci at points (2,11) and (2,1).
11. A hyperbolas with perpendicular asymptotes is called **perpendicular**. What does the equation of a perpendicular hyperbola look like?

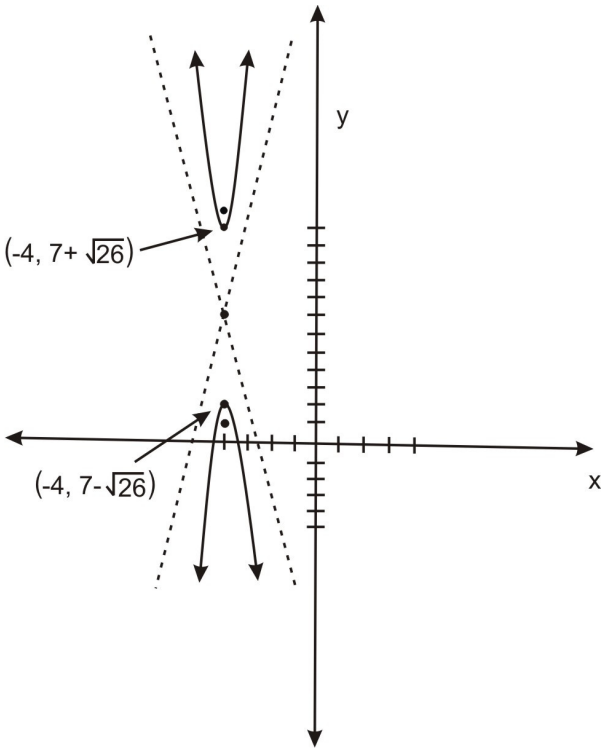
Review Answers

7. The distance between a point and a line is the shortest segment between the point and a point on the line. We have shown that *some* distance—not necessarily the shortest—between P and a point on the asymptote becomes infinitesimally smaller. This means that the shortest distance between P and the asymptote must also become shorter.

8.



9.



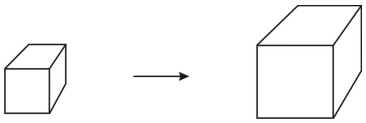
10. $\frac{(y-6)^2}{25} - \frac{(x-2)^2}{144} = 1$
11. The slopes of perpendicular lines are negative reciprocals of each other. This means that $\frac{a}{b} = \frac{b}{a}$, which, for positive a and b means $a = b$.

Applications

Vexing Questions from Ancient Greece

This application of hyperbolas is the reason Menaechmus, one of the first serious thinkers about conic sections, got into them in the first place.

It is thought that conic sections first attracted the attention of the Ancient Greeks when Menaechmus used two conics to find a solution to another famous Greek problem. A famous class of Greek problems involves constructing a length with the simplest tools possible. One of these problems is often referred to as the problem of "doubling a cube." The problem is to make a cube that is twice the volume of a given cube.



This, of course, does not simply mean making a cube whose sides are double in length. This would produce a cube with eight times the volume. If the first cube has side length one unit then it has volume $1^3 = 1$. In order for a cube to have double that volume, 2, it must have a side length of $\sqrt[3]{2}$. So the problem boils down to producing a segment whose length is $\sqrt[3]{2}$. Once this is done, the cube can be constructed out of segments of this length.

So Menaechmus was working on constructing a segment of length $\sqrt[3]{2}$. Now you are probably thinking, "Easy, just plug $\sqrt[3]{2}$ into your calculator." But Menaechmus didn't have a calculator into which he could type 2^(1/3). An even more important point is that *even if he did* have such a calculator, this wouldn't solve the problem in a way Menaechmus would have been pleased with. He was trying to use the *simplest* technology possible, and a calculator—even if he had one—wouldn't fit this bill. Plus, a calculator merely finds an estimate of this number, and the Greeks aimed to find a mathematical situation in which the *exact* number would be generated.

The very simplest tools the Greeks considered for making these kinds of constructions were an idealized compass and straightedge. Unknown to the Greeks at the time, a compass and straightedge don't do the job for this length, though they didn't know it at the time—it wasn't discovered until 1837, when Pierre Laurent Wantzel demonstrated that this construction is impossible.

According to some historians (Heath [footnote]), Menaechmus did find a solution to the problem and he used conic sections to do so. Before his time, the problem had already been reduced to finding a specific set of "mean proportionals", or a chain of numbers with equivalent subsequent ratios. In particular, to find numbers x and y such that 1 is to x as x is to y as y is to 2, or in fraction form:

$$\frac{1}{x} = \frac{x}{y} = \frac{y}{2}$$

How does this chain of proportions help find the number $\sqrt[3]{2}$? Well, looking at the first equality and cross-multiplying, we have $y = x^2$. Looking at the second equality, we have $y^2 = 2x$. Using the first equation to substitute x^2 for y in the second equation, we have $(x^2)^2 = 2x$, which reduces to $x^3 = 2$, the solution of which is $\sqrt[3]{2}$.

But solving the above chain of ratios is also equivalent to finding the intersection of the parabola $y = x^2$ and the hyperbola $xy = 2$. You'll examine why this is true in an upcoming review questions.

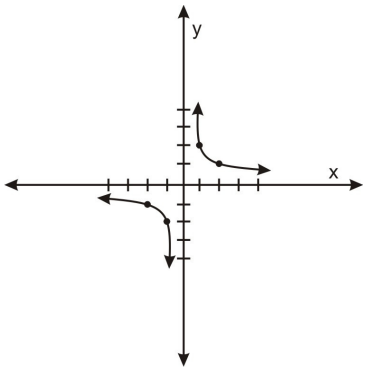
But wait! Hold the presses! Did I say the hyperbola $xy = 2$? But that doesn't look like the form of any of the hyperbolas we have looked at! This is because it is oriented with its foci on a diagonal, rather than horizontally or vertically. In the next section we will look at how rotating conic sections from their standard positions affects their forms. For now, you can check that $xy = 2$ indeed looks like a hyperbola by plugging in a few points near the origin—which you will do in the exercises below.

Review Questions

12. Draw a sketch of $xy = 2$ by plotting a few points near the origin.
13. What are the asymptotes of the hyperbola $xy = 2$? What are the foci?
14. Explain why solving the mean proportional $\frac{1}{x} = \frac{x}{y} = \frac{y}{2}$ is equivalent to finding the intersection of the parabola $y = x^2$ and the hyperbola $xy = 2$.
15. In the language of "compass and straightedge," though Menaechmus' construction can't be done with a traditional compass, it can be done with the generalized compass discussed in the previous section.

Explain how this tool could be used to "double" an existing cube.

Review Answers

12. 

13. The asymptotes are the $x-$ and $y-$axes. The foci are (2,2) and (-2,-2) (these are relatively hard to find, but it is relatively easy to show they are the foci once they are found.)
14. These two equations are obtained by looking at the first equality and cross multiplying, as well as setting the first term equal to the third term and cross multiplying. These two equalities hold exactly the same amount of information as the chain of equalities.
15. Using a compass and straightedge, a coordinate grid with unit lengths can be drawn. See http://en.wikipedia.org/wiki/Compass_and_straightedge_constructions for more on using a compass and straightedge to make geometric constructions. Upon this, the two shapes can be drawn using the generalized compass as discussed in the last session. The distance between the $y-axis$ and the intersection point is of length $\sqrt[3]{2}$.

Vocabulary

Hyperbola A conic section where the cutting plane intersects both sides of the cone, resulting in two infinite "U"-shapes curves.

Unbounded A shape which is so large that no circle, no matter how large, can enclose the shape.

Asymptote A line which a curve approaches as the curve and the line approach infinity, eventually becoming closer than any given positive number.

Perpendicular Hyperbola A hyperbola where the asymptotes are perpendicular.

1.5 General Algebraic Forms

Learning Objectives

- Understand how the cross sections of a cone relate to degree two polynomial equations.
- Understand what happens to the algebraic forms of conic sections when they are neither horizontally or vertically oriented.
- Recognize the algebraic form of different types of conic sections.

The Cross Sections of the Cone Are Degree Two Polynomial Equations

Let's examine all the equations of the conic sections we've studied in this chapter.

Ellipses: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where a and b are any positive numbers (the circle is the specific case when $a = b$).

Parabolas: $y = ax^2$ or $x = ay^2$ where a is any non-zero number.

Hyperbolas: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ or $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$, where a and b are any positive numbers.

All these equations have in common that they are degree-2 polynomials, meaning the highest exponent of any variable—or sum of exponents of products of variables—is two. So for example, here are some degree two polynomial equations in a more general form:

$$\begin{aligned} 2x^2 + 5y^2 - 3y + 4 &= 0 \\ x^2 - 3y^2 + x - y + 3 &= 0 \\ 10x^2 - y - 5 &= 0 \\ xy &= 2 \end{aligned}$$

Some of these probably already look like conic sections to you. For example, in the first equation, we can complete the square to remove the $-3y$ term and we will see that we have an ellipse. In the second equation we can complete the square twice to remove both the x and $-y$ terms and we will have a hyperbola. This is a hyperbola, not an ellipse, because the coefficient of the x^2 and y^2 terms have opposite signs.

The third equation is a parabola since there is an x^2 term and y term but not a y^2 term. Do you see how you can solve for y, putting the equation in the standard form for a vertically oriented parabola?

But what about the fourth equation? Like the others, it is a degree-2 polynomial, since the exponents of the x and y term sum to 2. But the fourth equation looks nothing like any of the forms for conic sections that we've examined so far. Nonetheless, as we saw in the last section, $xy = 2$ appears to be a hyperbola with foci (2,2) and (-2,-2). The reason it doesn't fit either of the standard forms for hyperbolas is because it is *diagonally* oriented, rather than horizontally or vertically oriented (do you see how its two foci lie on a diagonal line, rather than a horizontal or vertical line?)

In order to see how such differently-oriented conic sections fit into our standard forms, we need to rotate them so that they are either horizontally or vertically oriented.

Rotation of Conics

Remember in the section on parabolas we discussed rotating objects in the plane. In that section we showed that if we take a point $P = (x, y)$ and rotate it θ degrees, it changes the $x-$coordinate to $x' \cos(\theta) - y' \sin(\theta)$ and the $y-$coordinate changes to $x' \sin(\theta) + y' \cos(\theta)$. (Note: in that section, because we were looking at a plane embedded in space, we happened to be looking at the $xz-$plane. We also we're using the "prime" symbols x' and y'. Also, in the particular case of that section we were rotating the cone clockwise. It is more standard to rotate counter-clockwise, so the signs on the sine functions have changed. So you can simply substitute x' for x and y' for z into the formulas in that section, switch the signs on the sine functions, and come up with these rotation rules for the $xy-$plane.)

Now suppose we have an equation of degree two polynomials, such as the $xy = 2$ example discussed above. In order to put it into the more recognizable form of a ellipse, parabola, or a hyperbola, we need to rotate it in such a way so that rotated version has no xy term. So we need to find an appropriate angle θ such that changing the $x-$coordinate to $x' \cos(\theta) - y' \sin(\theta)$ and the $y-$coordinate to $x' \sin(\theta) + y' \cos(\theta)$ results in

an equation with no xy term. We need to investigate what happens to a degree-two polynomial equation. Such an equation can be written in the form:

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$$

If $C = 0$, such as in the first three examples at the beginning of this section, we are done, as there is no xy term and we already know how to classify these into conics. If $C \neq 0$, we need to rotate the curve so that we have an equation with no xy term. When x is replaced by $x' \cos(\theta) - y' \sin(\theta)$ and y is replaced by $x' \sin(\theta) + y' \cos(\theta)$, only the first three terms of this equation are in danger of producing an xy term. To see if we can determine if an appropriate angle θ can always be found, let's substitute our new variables in for x and y into the first three terms of the equation:

$$A(x' \cos(\theta) - y' \sin(\theta))^2 + B(x' \sin(\theta) + y' \cos(\theta))^2 + C(x' \cos(\theta) - y' \sin(\theta))(x' \sin(\theta) + y' \cos(\theta))$$

Then, let's multiply this expression out, but only examine the terms that are a multiple of $x'y'$, since that is what we're trying to eliminate.

$$-2Ax'y' \cos(\theta) \sin(\theta) + 2Bx'y' \cos(\theta) \sin(\theta) + Cx'y'(\cos^2(\theta) - \sin^2(\theta))$$

This reduces to:

$$(-2A \cos(\theta) \sin(\theta) + 2B \cos(\theta) \sin(\theta)) + C(\cos^2(\theta) - \sin^2(\theta)))x'y'$$

So we are interested in whether or not there is an angle θ such that the above coefficient of $x'y'$ is zero:

$$\begin{aligned}(-2A \cos(\theta) \sin(\theta) + 2B \cos(\theta) \sin(\theta)) + C(\cos^2(\theta) - \sin^2(\theta)) &= 0 \\ C(\cos^2(\theta) - \sin^2(\theta)) &= (2A \cos(\theta) \sin(\theta) - 2B \cos(\theta) \sin(\theta)) \\ C(\cos^2(\theta) - \sin^2(\theta)) &= 2(A - B)(\cos(\theta) \sin(\theta))\end{aligned}$$

Separating the θ terms from the A, B, and C terms, we have:

$$\frac{2 \cos(\theta) \sin(\theta)}{\cos^2(\theta) - \sin^2(\theta)} = \frac{C}{A - B}$$

But remember the double angle formulas $\sin(2\theta) = 2 \cos(\theta) \sin(\theta)$ and $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ (see http://authors.ck12.org/wiki/index.php/Trigonometric_Identities#Double-Angle_Identities for more on the double angles). This means that we need to find an angle such that:

$$\frac{\sin(2\theta)}{\cos(2\theta)} = \frac{C}{A - B}$$

or:

$$\tan(2\theta) = \frac{C}{A - B}$$

Can we find such an angle? Well remember the graph of tangent (see http://authors.ck12.org/wiki/index.php/Circular_Functions#y_.3D_tan.28x.29.2C_The_Tangent_Graph for more on the tangent graph). It spans all values, so no matter what C, A, and B equal, we can find the appropriate tangent value. The only time this expression gives us trouble is if the denominator is zero, or $A = B$. But in this case there's an easy solution that you will find in the exercise below.

For the other cases, you can use the inverse tan function on your calculator or computer program to find out the value of 2θ, and then divide by 2 to find the value of θ.

Example: Rotate the following conic section so that it is oriented either horizontally or vertically, and then analyze the result:

$$79x^2 + 37y^2 + 42\sqrt{3}xy + \left(-200\sqrt{3} - 8\right)x + \left(8\sqrt{3} - 200\right)y + 4 = 0$$

Solution: To find how much to rotate this conic we need to solve for θ in the equation $\tan(2\theta) = \frac{C}{A-B}$. We have $C = 42\sqrt{3}, A = 79$, and $B = 37$, so we have:

$$\tan(2\theta) = \frac{42\sqrt{3}}{79 - 37}$$
$$\tan(2\theta) = \sqrt{3}$$

The angle that solves this equation is $\theta = \frac{\pi}{6}$ (or $30°$).

So we replace x with $x'\cos(30) - y'\sin(30) = \frac{\sqrt{3}}{2}x' - \frac{1}{2}y'$, and y with $x'\sin(30) + y'\cos(30) = \frac{1}{2}x' + \frac{\sqrt{3}}{2}y'$, giving us:

$$79\left(\frac{\sqrt{3}}{2}x' - \frac{1}{2}y'\right)^2 + 37\left(\frac{1}{2}x' + \frac{\sqrt{3}}{2}y'\right)^2 + 42\sqrt{3}\left(\frac{\sqrt{3}}{2}x' - \frac{1}{2}y'\right)\left(\frac{1}{2}x' + \frac{\sqrt{3}}{2}y'\right) +$$
$$\left(-200\sqrt{3} - 8\right)\left(\frac{\sqrt{3}}{2}x' - \frac{1}{2}y'\right) + \left(8\sqrt{3} - 200\right)\left(\frac{1}{2}x' + \frac{\sqrt{3}}{2}y'\right) + 4 = 0$$

Multiplying through by 4 we have:

$$79\left(\sqrt{3}x' - y'\right)^2 + 37\left(x' + \sqrt{3}y'\right)^2 + 42\sqrt{3}\left(\sqrt{3}x' - y'\right)\left(x' + \sqrt{3}y'\right) +$$
$$\left(-200\sqrt{3} - 8\right)\left(2\sqrt{3}x' - 2y'\right) + \left(8\sqrt{3} - 200\right)\left(2x' + 2\sqrt{3}y'\right) + 16 = 0$$

Multiplying out we have:

$$237(x')^2 - 158\sqrt{3}x'y' + 79(y')^2 + 37(x')^2 + 74\sqrt{3}x'y' + 111(y')^2 + 126(x')^2 + 84\sqrt{3}x'y' - 126(y')^2$$
$$- 1200x' - 16\sqrt{3}x' + 400\sqrt{3}y' + 16y' + 16\sqrt{3}x' - 400x' + 48y' - 400\sqrt{3}y' + 16 = 0$$

Which simplifies to:

$$400(x')^2 + 64(y')^2 - 1600x' + 64y' + 16 = 0$$

Which, divided by 16 is:

$$25(x')^2 + 4(y')^2 - 100x' + 4y' + 1 = 0$$

Grouping the x' and y' terms and completing the squares, we have:

$$25((x')^2 - 4x' + 4) - 100 + 4\left((y')^2 + y' + \frac{1}{4}\right) - 1 + 1 = 0$$
$$25((x')^2 - 2)^2 + 4\left((y')^2 + \frac{1}{2}\right)^2 = 100$$
$$\frac{((x')^2 - 2)^2}{4} + \frac{\left((y')^2 + \frac{1}{2}\right)^2}{25} = 1$$

We recognize this as an ellipse, centered at the point $\left(2, -\frac{1}{2}\right)$.

Review Questions

1. What's the easy solution to the above problem when $A = B$?
2. Use the above answer or a calculator or computer program to determine how much you would have to rotate the following conic to eliminate the xy term and orient it either horizontally or vertically: $x^2 + y^2 + 6x - 3y + 10xy + 100 = 0$.
3. Use the above answer or a calculator or computer program to determine how much you would have to rotate the following conic to eliminate the xy term and orient it either horizontally or vertically: $5x^2 - 3x + 4xy + 21 = 0$.
4. Rotate the following conic section so that it is oriented either horizontally or vertically, and then analyze the result: $3x^2 + 3y^2 + \left(4\sqrt{2} + 2\right)xy + 4\sqrt{2}x + 4\sqrt{2}y - 4 = 0$

Review Answers

1. $\theta = \frac{\pi}{2}$
2. $\theta = \frac{\pi}{2}$
3. $\theta \approx 31.72°$
4. To figure out the angle of rotation, since $A = B$ we have $\theta = \frac{\pi}{4}$ (or 45°). After shifting the equation by this amount, we have a relation, $4x^2 + 8x - y^2 + 8y - 4 = 0$. Completing the square, this results in a shifted hyperbola: $\frac{(x+1)^2}{1} - \frac{(y-2)^2}{4} = 1$.

General Algebraic Forms

How can we look at a degree-2 polynomial equation and determine which conic section it depicts?

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$$

When $C = 0$ we have already discussed how to determine which conic section the equation refers to. In summary, if A and B are both positive, the conic section is an ellipse. This is also true of A and B are both negative, as the entire equation can be multiplied by -1 without changing the solution set. If A and B differ in sign, the equation is a hyperbola, and if A or B equals zero the equation is a parabola.

There are a few new, more general, rules I will show you that give more information about the case when $C \neq 0$ and hence the conic section needs to be rotated to achieve horizontal or vertical orientation.

If $C^2 < 4AB$, the equation is an ellipse (note when $C = 0$ this holds whenever A and B are the same sign, which is consistent with our simpler rule stated above.)

If $C^2 > 4AB$, the equation is an hyperbola (note when $C = 0$ this holds whenever A and B are the opposite sign, which is consistent with our simpler rule stated above.)

If $C^2 = 4AB$, the equation is a parabola (note when $C = 0$, either A or B equals zero, which is consistent with our simpler rule stated above.)

Review Questions

5. State what type of conic section is represented by the following equation: $5x^2 + 6y^2 + 2x - 5y + xy = 0$.
6. State what type of conic section is represented by the following equation: $x^2 + 3y - 20xy + 20 = 0$.
7. The rules above do not account for "degenerate" conic sections, that is the conic section that looks like an X made by the intersection of a plane containing the line at the center of the cone. Explain the conditions on the coefficients that lead to the degenerate conic sections.

Review Answers

5. Ellipse
6. Hyperbola
7. $C = 0$, and either $A = 0$ or $B = 0$ (or both).

Image Sources

(1) http://upload.wikimedia.org/wikipedia/commons/d/d8/Focke_Garten_0004.jpg. Creative Commons Attribution 2.5.

(2) http://commons.wikimedia.org/wiki/File:Leonardo_parabolic_compass.JPG. Public Domain.

(3) http://commons.wikimedia.org/wiki/File:Parabolic_reflection_1.svg. GNU-FDL.

(4) http://commons.wikimedia.org/wiki/File:Cassegrain_telescope.png. GNU-FDL.

(5) http://commons.wikimedia.org/wiki/File:DSN_Antenna_details.jpg. Public Domain.

(6) http://commons.wikimedia.org/wiki/File:Headlight_lens_optics_schematic.png. GNU-FDL.

(7) http://upload.wikimedia.org/wikipedia/commons/8/87/Kepler-first-law.png. CC by SA 3.0.

Chapter 2

Analyzing Functions

2.1 Identifying Functions

Learning objectives

- Determine if a relation is a function or not
- Determine the domain and range of a function
- Graph, or read a graph of basic real functions on subintervals of their domains
- Determine the average rate of change of a function

Introduction

Consider two situations shown in the boxes below:

Table 2.1:

Situation 1: You are selling candy bars for a school fundraiser. Each candy bar costs $3.00	Situation 2: You collect data from several students in your class on their ages and their heights: (18,65"), (17,64"), (18,67"), (18,68"), (17,66")

In the first situation, let the variable x represent the number of candy bars that you sell, and let y represent the amount of money you make. If you sell x candy bars, you will make $y = 3x$ dollars. For example, if you sell 25 candy bars, you will make $3(25) = \$75.00$. Notice that you can use the number of candy bars you sell to "predict" how much money you will make.

Now consider the second situation. Can you use the data to predict height, given age?

This is not the case in the second situation. For example, if a student is 18 years old, there are several heights that the student could be.

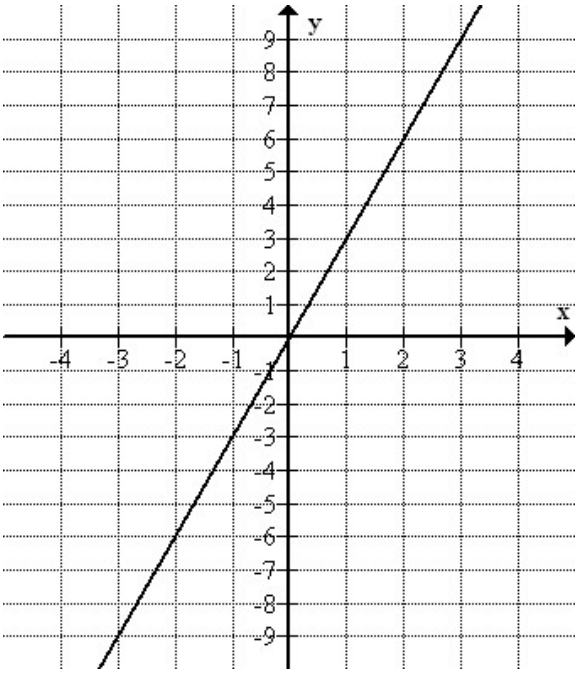
The first situation is an example of a *function,* and the second example is *not* a function. In this lesson you will learn to identify mathematical relations that are functions and to describe them in particular ways. This lesson will prepare you to learn about several kinds of functions throughout this chapter.

Relations and Functions

Both situations above are relations. A **relation** is simply a relationship between two sets of numbers or data. For example, in the second situation, we created a relationship between students' ages and heights, just by writing each student's information as an ordered pair. In the first situation, there is a relationship between the number of candy bars you sell and the amount of money you make. The first example is different from the second because it represents a **function**: *every* x *is paired with only one y.*

We can represent functions in many ways. Some of the most common ways to represent functions include: sets of ordered pairs, equations, and graphs. The figure below shows the same function depicted in each of these representations:

Table 2.2:

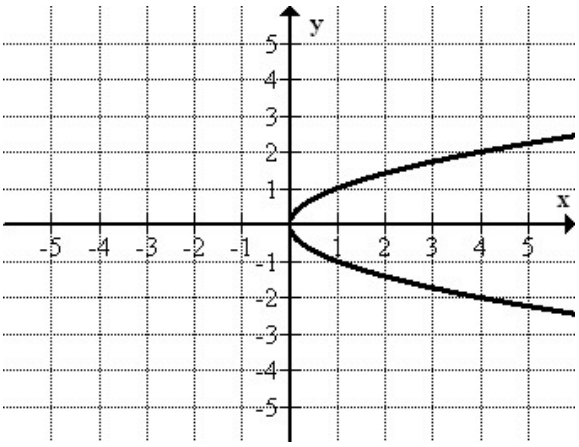
Representation	**Example**
Set of ordered pairs	(1,3), (2,6), (3,9), (4,12) (a subset of the ordered pairs for this function)
Equation	$y = 3x$
Graph	

In contrast, the relation shown in the figure below is *not* a function:

Table 2.3:

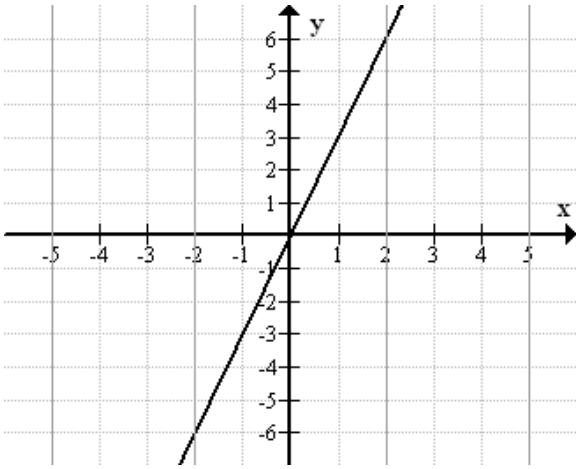
Representation	**Example**
Set of ordered pairs	(4, 2), (4, -2), (9, 3), (9, -3) (A subset of the ordered pairs for this relation)
Equation	$x = y^2$

Table 2.3: (continued)

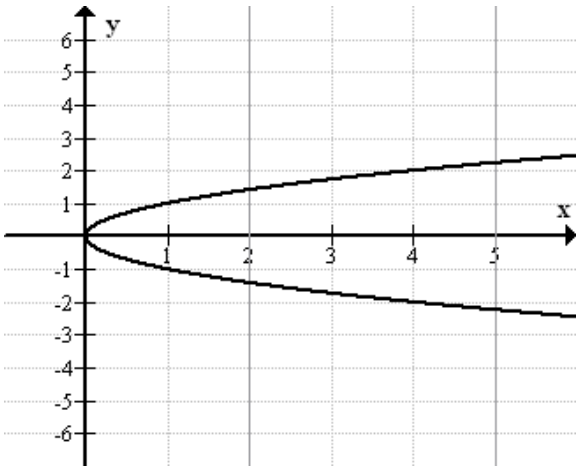
Representation	Example
Graph	

To verify that this relation is not a function, we must show that at least one x value is paired with more than one y value. If you look at the first representation, the set of ordered pairs, you can see that 4 is paired with 2 and with -2. Similarly, 9 is paired with 3 and with -3. Therefore the relation is not a function. It is harder to identify a function from an equation, unless you are familiar with different kinds of equations. (If you are not already familiar with different kinds of functions and their graphs, you will be by the end of this chapter!) If we look at the graph above, we can see that, except for $x = 0$, the x values of the relation are each paired with two y values. Therefore the relation is not a function.

One way to determine whether a relation is a function when looking at a graph is by doing a "vertical line test". If a vertical line can be drawn anywhere on the graph such that the line crosses the relation in two places, then the relation is not a function. If all possible vertical lines will only cross the relation in one place, then the relation is a function. This works because if a vertical line crosses a relation in more than one place it means that there must be two y values corresponding to one x value in that relation. For example, the graph below of $y=3x$ shows it is a function because any vertical line that is drawn only crosses the relation in one place.

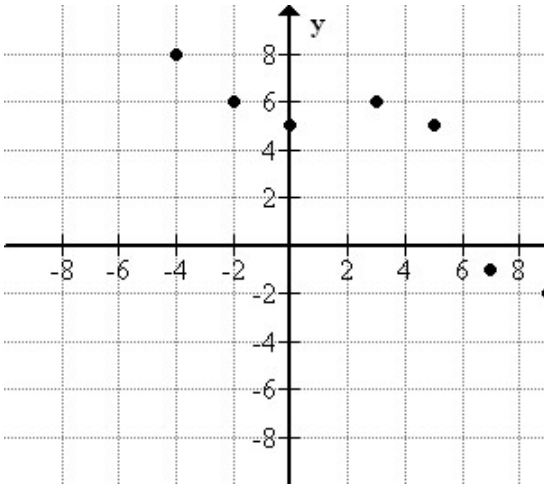


Conversely, the graph below of $x=y^2$ shows it is not a function because a vertical line can be drawn that crosses the relation in two places.



Example 1: Determine if the relation is a function or not

Table 2.4:

a. (2, 4), (3, 9), (5, 11), (5, 12)	b. 

Solution:

a. (2, 4), (3, 9), (5, 11), (5, 12)

This relation is not a function because 5 is paired with 11 and with 12.

b. This relation is a function because every x is paired with only one y.

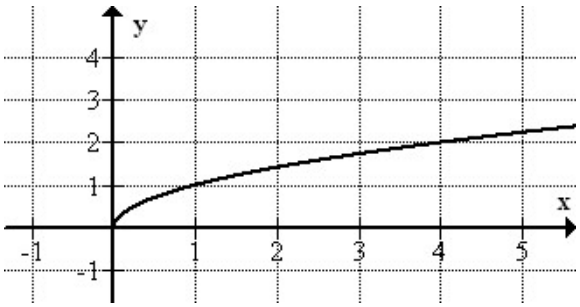
Once you are able to determine if a relation is a function, you should then be able to state the set of x values and the set of y values of the function. Next we will define these sets as the **domain** and the **range** of a function, and we will look at some specific examples of functions and their domains and ranges.

Independent variable, domain

The **domain** of a function is defined as the set of all x values for which the function is defined. For example, the domain of the function $y = 3x$ is the set of all real numbers, often written as $\mathbb{R}$. This means that x can be any real number. Other functions have restricted domains. For example, the domain of the

function $y = \sqrt{x}$ is the set of all real numbers greater than or equal to zero. The domain of this function is restricted in this way because the square root of a negative number is not a real number. Therefore, the domain is restricted to non-negative values of x so that the function values will be defined.

It is often easy to determine the domain of a function by (1) considering what restrictions there might be and (2) looking at a graph. For example, we can see that the function $y = \sqrt{x}$, shown in the graph below, has a domain of all real numbers greater than or equal to zero because the graph only exists for x values that are greater than or equal to zero.



Example 2: State the domain of each function:

Table 2.5:

a. $y = x^2$	b. $\frac{1}{x}$	c. (2, 4), (3, 9), (5, 11)

Solution:

a. $y = x^2$

The domain of this function is the set of all real numbers. There are no restrictions.

b. $\frac{1}{x}$

The domain of this function is the set of all real numbers except $x = 0$. The domain is restricted this way because a fraction with denominator zero is undefined.

c. (2, 4), (3, 9), (5, 11)

The domain of this function is the set of x values: $\{2, 3, 5\}$.

The variable x is often referred to as the **independent** variable, while the variable y is referred to as the **dependent** variable. We talk about x and y this way because the y values of a function depend on what the x values are. That is why we also say that "y is a function of x." For example, the value of y in the function $y=3x$ depends on what x value we are considering. If $x = 4$, we can easily determine that $y = 3(4) = 12$. Returning to the situation in the introduction, we can say that the amount of money you take in *depends* on the number of candy bars you sell.

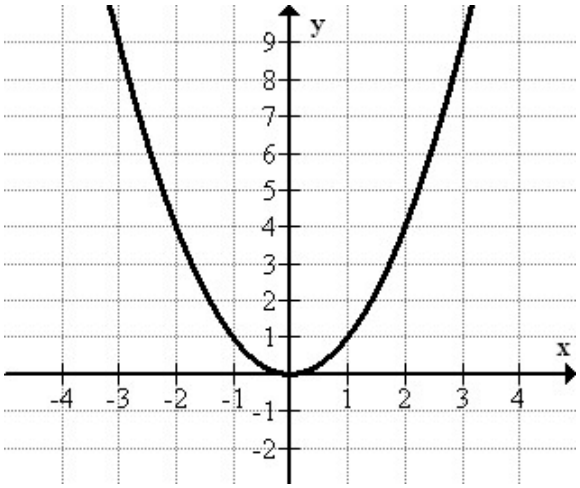
When we are working with a function in the form of an equation, there is a special notation we can use to emphasize the fact that y is a function of x. For example, the equation $y = 3x$ can also be written as $f(x) = 3x$. It is important to remember that $f(x)$ represents the y values, or the function values, and that the letter f is *not* a variable. (That is, $f(x)$ does *not* mean that we are multiplying.)

Now that we have considered the domain of a function, we will turn to the range.

Dependent Variable, Range

The **range** of a function is defined as the set of all y values for which a function is defined. Just as we did with domain, we can examine a function and determine its range. Again, it is often helpful to think about what restrictions there might be, and what the graph of the function looks like. Consider for example the function $y = x^2$. The domain of this function is $\mathbb{R}$, all real numbers, but what about the range?

The range of the function is the set of all real numbers greater than or equal to zero. This is the case because every y value is the square of an x value. If we square positive and negative numbers, the result will always be positive. If x = 0, then y = 0. We can also see the range if we look at a graph of $y = x^2$. Notice below that the y values are all greater than or equal to zero.



Example 3: State the domain and range of the function $y = \frac{2}{x}$

Solution: The domain and range of the function $y = \frac{2}{x}$

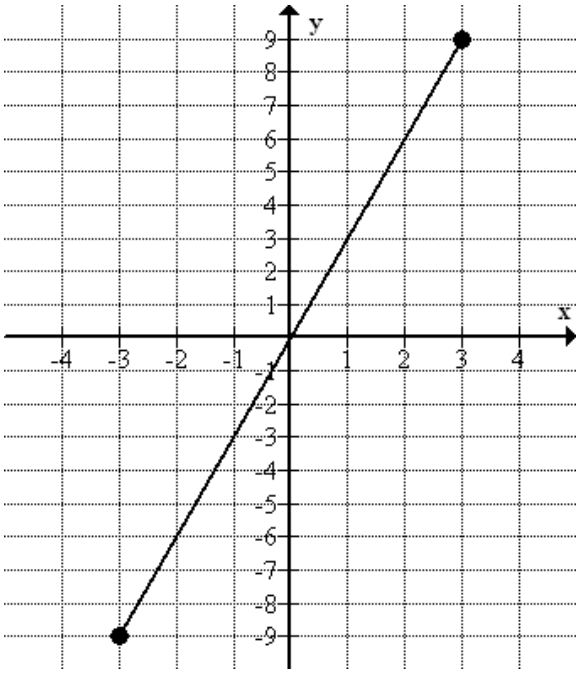
For this function, we can choose any x value except $x = 0$. Therefore the domain of the function is the set of all real numbers except $x = 0$.

The range is also restricted to the non-zero real numbers, but for a different reason. Because the numerator of the fraction is 2, the numerator can never equal zero, so the fraction can never equal zero.

Now that we have defined what it means for a relation to be a function, and we have defined domain and range of a function, we can look at some specific examples of functions and their graphs.

Real values, intervals

A function is defined as a **real function** if both the domain and the range are sets of real numbers. Many of the functions you have likely encountered before are real functions, and many of these functions have Domain = $\mathbb{R}$. Consider, for example, the function $y = 3x$. The graph of this function, shown at the beginning of this lesson, is shown again below, slightly varied.



You may already be familiar with the graphs of lines. In particular, you may already be in the habit of placing arrows at the ends. We do this in order to indicate that the line will continue forever in both the positive and negative directions, both in terms of the domain and the range. The line above, however, only shows the function $y = 3x$ on the interval [-3, 3]. The square brackets indicate that the graph includes the endpoints of the interval $x = -3$ and $x = 3$. We call this a **closed interval**. A closed interval contains its endpoints. In contrast, an **open interval** does *not* contain its endpoints. We indicate an open interval with parentheses. For example, (-3, 3) indicates the set of numbers between -3 and 3, *not including* -3 and 3. You may have noticed that the open interval notation looks like the notation for a point (x, y) in the plane. It is important to read an example or a homework problem carefully to avoid confusing a point with an interval!

When you are graphing functions in general, especially those for which the domain is the set of all real numbers ($D = \mathbb{R}$), you will necessarily be showing a subset of the domain. In studying functions, you may also be asked to describe what the graph of a function looks like on a particular subset of its domain. The table below summarizes the kinds of intervals you may need to consider while studying functions and their domains:

Table 2.6:

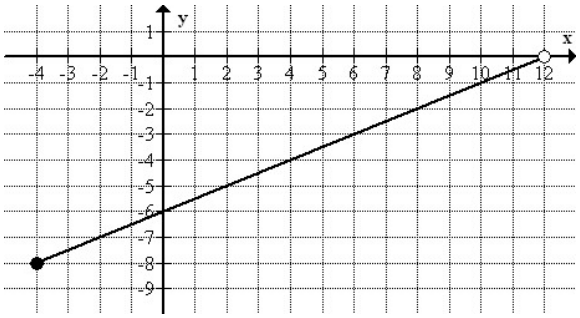
Interval notation	Inequality notation	Description
$[a, b]$	$a \leq x \leq b$	The value of x is between a and b, including a and b, where a, b are real numbers.
(a, b)	$a < x < b$	The value of x is between a and b, *not* including a and b.
$[a, b)$	$a \leq x < b$	The value of x is between a and b, including a, but not including b.
$(a, b]$	$a < x \leq b$	The value of x is between a and b, including b, but not including a.

Table 2.6: (continued)

Interval notation	Inequality notation	Description
(a, ∞)	$x > a$	The value of x is strictly greater than a.
$[a, \infty)$	$x \geq a$	The value of x is greater than or equal to a
$(-\infty, a)$	$x < a$	The value of x is strictly less than a
$(-\infty, a]$	$x \leq a$	The value of x is less than or equal to a.

Example 4: Sketch the graph of the function $f(x) = \frac{1}{2}x - 6$ on the interval [-4, 12)

Solution: The figure below shows a graph of $f(x) = \frac{1}{2}x - 6$ on the given interval:

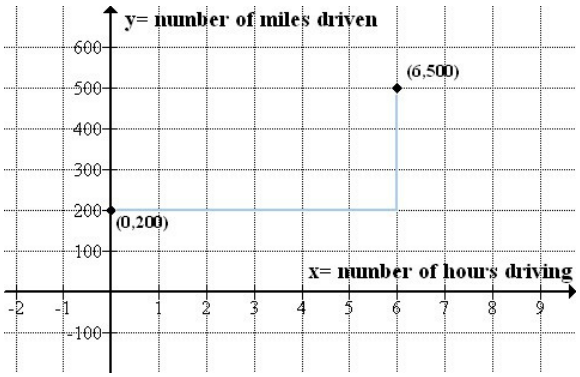


Above we have considered several different kinds of functions. If you have looked at the graphs of functions such as $f(x) = \frac{1}{2}x - 6$ and $y = x^2$, you may have noticed that the shapes of the graphs are quite different. Later in the chapter you will study different "families" of functions and their graphs. Here we will consider one aspect of functions that requires us to consider specific intervals of the function.

Average rate of change

Consider the following situation: you are on a week long road trip with your friend. When you begin to drive on the second day, you have already driven a total of 200 miles. After 6 hours of driving on the second day, you have driven a total of 500 miles. On average, how many miles did you drive per hour on the second day of the trip?

The graph below shows this situation, with the x axis representing the number of hours driving (on the second day), and the y axis representing the number of miles driven. The first point on the graph, (0,200), says that at the beginning of the second day you have already driven 200 miles. The second point on the graph, (6,500), says that after 6 hours of driving on the second day you have driven 500 miles total.



Notice that in total, during your 6 hours of driving, you driven 300 miles. The rate at which you drove is 300 miles in 6 hours, or 50 miles per hour. We refer to this rate as the **average rate of change** because it is an average across the 6 hours. That is, you did not necessarily drive 50 miles every hour. There could have been one hour where you drove 70 miles and another hour where you drove only 30 miles.

We can represent the average rate of change on the graph by indicating how much each quantity has changed: The y values increased by 300, and the x values increased by 6. The average rate of change is the ratio of these changes in each variable. This is how we can define average rate of change in general:

$$\text{Average rate of change } = \frac{\text{change in y}}{\text{change in x}}$$

We can examine the average rate of change of a function, whether it is represented as data, as in the previous example, or by an equation.

Example 5: Find the average rate of change of each function on the given interval

Table 2.7:

a. $f(x) = x^2$ on $[0, 2]$	b. $f(x) = 4x$ on $[1, 7]$

Solution:

a. $f(x) = x^2$

The endpoints of the interval are (0,0) and (2,4). Therefore the change in y is 4 and the change in x is 2. The average rate of change is $4/2 = 2$.

b. $f(x) = 4x$

The endpoints of the interval are (1,4) and (7,28). Therefore the change in y is 28 - 4 = 24 and the change in x is 7 - 1 = 6. The average rate of change is $24/6 = 4$.

Notice that the average rate of change of the function $f(x) = 4x$ is the slope of the line, 4. While a linear function has a constant slope, other functions, such as $f(x) = x^2$, will not. You will explore this idea in greater detail in your study of calculus.

Lesson Summary

In this lesson we have defined what it means for a relation between two variables (such as x and y) to be a function, and we have considered several key aspects of functions: the domain and range of a function, sub-intervals of the domain of a real function, and the average rate of change of a function. We have

explored these ideas in the context of specific examples of linear and non-linear functions. In the coming lessons you will rely on the ideas introduced here to further develop your understanding of functions.

Points to Consider

1. How can you determine if a relation is a function?
2. What kinds of functions have domain and range both equal to the set of all real numbers?
3. How is average rate of change different for linear and non-linear functions?

Review Questions

1. Determine if each relation is a function:

Table 2.8:

a. (-1,4), (0, 3), (1, 5), (1, 7), (2, 15)	b. $y = x$

2.
3. State the domain and range of each relation in question 1.
4. Give an example of a function for which the domain and range are equivalent to each other.
5. Write each interval using interval notation.

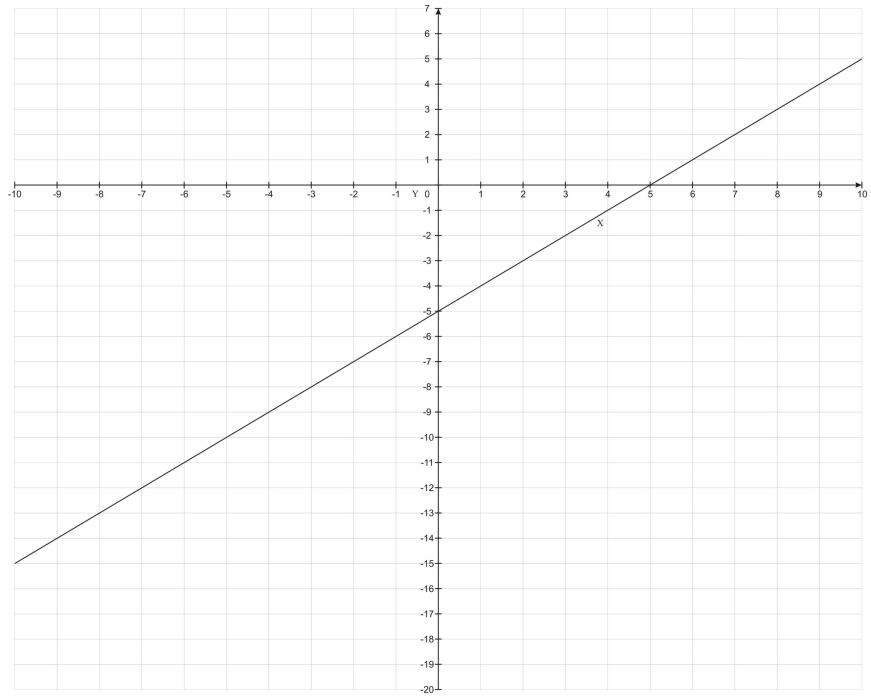
Table 2.9:

a. $x>5$	b. -4 $x<7$

6.
7. Sketch a graph of the function $f(x) = x - 5$ on the interval [-10, 10]
8. What is the average rate of change of the function $f(x) = 3x^2$ on the interval [2, 5]?
9. Consider the function $f(x) = 2x - 7$. a. Choose two points on the graph of $f(x)$ and calculate the average rate of change.
   b. Choose two other points and calculate the average rate of change.
   c. Compare your answers in parts a and b.
10. Consider the function $f(x) = ax^4$a. If $a > 0$, what is the range of this function?
    b. If $a < 0$, what is the range of this function?
11. Give an example of a relation that is not a function, and explain why it is not a function.
12. Consider the function $f(x) = ax^4$ from question 8. Assume that $a > 0$. a. What is the average rate of change of the function on the interval [0, 1]?
    b. What is the average rate of change of the function on the interval [0, 0.1]?

Review Answers

1. a. Not a function b. Is a function
2. a. D: {-1, 0, 1, 2}; R: {3, 4, 5, 7, 15} b. D:$\mathbb{R}$; R: $\mathbb{R}$
3. Answers will vary. The function in 1b is an example.
4. a. $(5, \infty)$b. [-4,7)

5. 

6. The two points are (2, 12) and (5, 75). The average rate of change is $63/3 = 21$.
7. a. and b. Choice of points will vary, but the average rate of change for both a and b should be 2. c. They should be equal to 2, the slope of the function, because this is a line.
8. a. If $a > 0$, R: y 0b. If $a < 0$, R: y 0
9. Answers will vary. Example: (2, 3) (2, 4) (4, 5) is not a function because the element 2 in the domain is paired with more than one element of the range.
10. a. ab. $\frac{.0001a}{.1} = 10^{-3}a$

2.2 Minimums and Maximums

Learning objectives

- Analyze a situation and determine if it is maximization or minimization
- Write an expression to represent such a situation
- Use a graphing utility to find a maximum or minimum

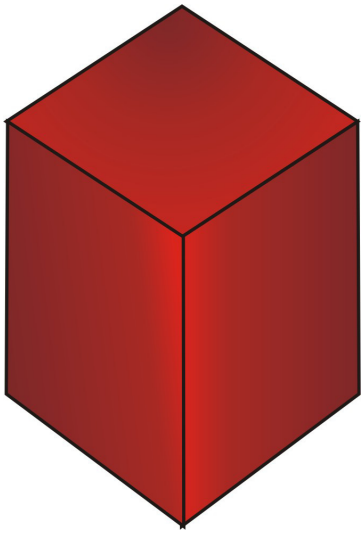
Introduction

Consider the following example: you run a business making boxes of cookies. While you earn money by selling cookies, it costs you money to construct every box. (These costs include the materials, as well as all of the costs to run the business, such as paying your employees!)

As the owner of the business, you will want to **maximize** your profits, and you will want to **minimize** your costs. That is, you want to make the most money that you can, while spending as little as you can in the process. In this chapter you will examine situations in which you will want to maximize or minimize some quantity. You will learn how to identify maxima and minima by writing functions to represent situations, and by examining the graphs of these functions.

Min-Max situations

In the example in the introduction, one quantity would be maximized, while the other would be minimized. In general, however, not every situation would necessarily include both of these ideas. For example, consider a situation in which you want to make a box that has a volume of 12 cm^3, but you want the surface area of the box to be as small as possible.



In this situation, you want to find the minimum surface area, given a fixed volume.

You might also have a situation in which you are given a fixed amount of material with which to make the box, and you want to make as *large* a box as possible. In this situation, you would want to find the *maximum* volume of the box, given a fixed surface area.

Whether you will be looking for a maximum or a minimum depends on the specific situation. Consider the situations in example 1.

Example 1: In each situation determine if a quantity should be maximized or minimized.

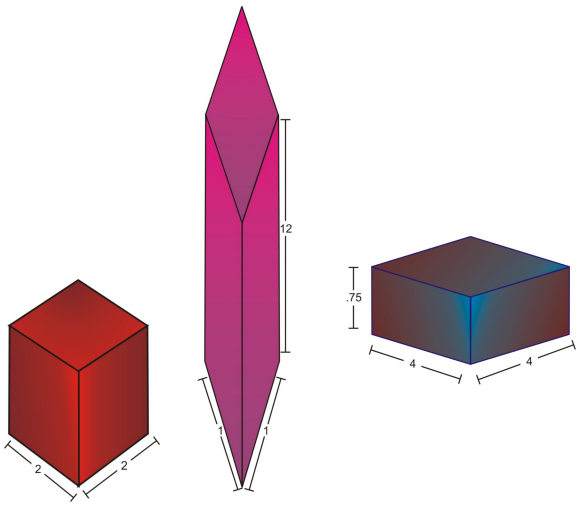
a. You have 100 feet of fence to enclose a field, and you want to create the largest field possible.
b. You run a factory that packages toilet paper, and you want to use the least amount of plastic possible for each roll.

Solution:

a. This situation involves maximizing the area of the field.
b. This situation involves minimizing the amount of plastic used per roll. (This would be the surface area of a cylinder.)

Expressions to be minimized or maximized

Consider the example above of the box with volume 12 cm^3. The volume of a box is the product of its three dimensions: length, width, and height. So we can express the volume of the box as an equation: $LWH = 12$. There are infinitely many boxes with volume 12. Three possible boxes are shown below:



We can find the surface area of each box by adding up the area of the six faces. The first box has four faces with area 6 and 2 faces with area 4, so the surface area of the box is 32 cm^2. Similarly, you can find that the second box has surface area 50cm^2. Finally the surface area of the third box is 44cm^2. Therefore, of these three boxes, the first one has the smallest surface area. However, this box does not necessarily have the smallest surface area possible for a box with the given volume. To find the box with minimum surface area requires that we know a bit more information about the box.

Let's consider a slightly simpler situation: a box with a square base and a fixed volume of 12 cm^3. Now let the length and the width of the box be x cm, and the height be h cm. We can write the volume equation as $x \cdot x \cdot h = x^2h = 12$. We can also express the surface area in terms of x and h:

$$\text{Surface area} = S = 4xh + 2x^2$$

(Make sure you understand this formula: The base and the top are squares with area $= x^2$ and the four sides are each rectangles of area equal to xh).

We can express the surface area as a function of x if we consider the volume equation and the surface area equation as a system of equations:

$$\begin{cases} x^2h = 12 \\ 4xh + 2x^2 = S \end{cases}$$

We want to work with the surface area equation since that is what we want to minimize. However, there are currently three variables in that equation: S, x, and h. It will be easier to graph and analyze surface area if we can express S in terms of just one other variable. So, we want to use substitution to get rid of one of the variables. We can use the volume equation to rewrite the surface area equation as a function of x.

First, rewrite the volume equation:

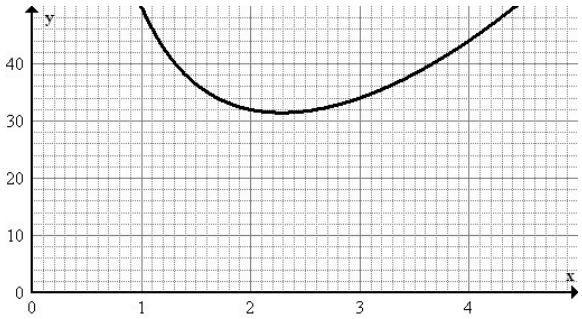
$x^2h = 12 \Rightarrow h = \frac{12}{x^2}$

Now, use substitution:

$$\begin{aligned} S(x) &= 4xh + 2x^2 \\ &= 4x\left(\frac{12}{x^2}\right) + 2x^2 \\ &= \frac{48}{x} + 2x^2 \end{aligned}$$

The values of the function $S(x)$ represent different possibilities for the surface area of the box, given that the base is a square, and given that the volume of the box is 12 cm^3. Therefore to identify the minimum surface area, we need to find the lowest function values for $S(x)$. When you study calculus, you will learn algebraic techniques for determining maximum and minimum values of a function. Here, we will rely on our knowledge of graphing to identify the minimum value of $S(x)$.

The graph below shows the function $S(x)$ on the interval [0,5]. It is important to note that the function $S(x)$ has the domain of all real numbers except zero, but we are only interested in x values greater than zero since x represents the side length of the box.

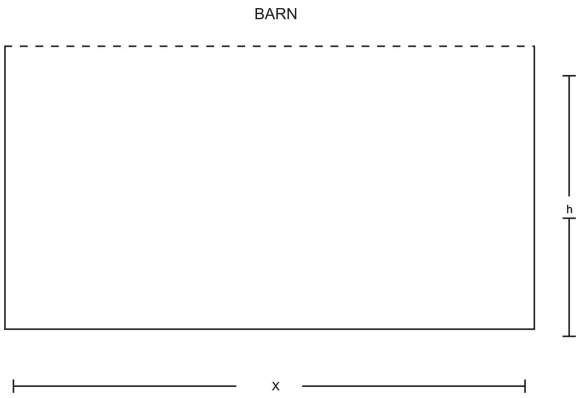


By examining the graph, we can see that the lowest point is between $x = 2$ and $x = 3$. If you use a "minimum" function on a graphing utility, you will find that the minimum point is approximately (2.3, 31.4). This tells us that when the side length of the box is approximately 2.3 cm, the surface area is approximately 31.4 cm^2, which is the smallest it can be.

The example below asks you to find a maximum value.

Example 2: You have 100 feet of fence with which to enclose a plot of land on the side of a barn. You want the enclosed land to be a rectangle. What size rectangle should you make with the fence in order to maximize the area of the rectangular enclosure?

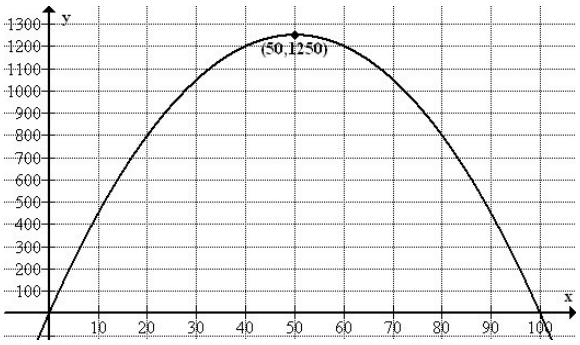
Solution: The plot of land will look like the picture below:



The area of the rectangular plot is the product of its length and width. We can write the area as a function of x: $A(x) = xh$. We can eliminate h from the equation if we consider that we have 100 feet of fence, and we write an equation about how we are using that 100 feet of fence: $x + 2h = 100$. (The fourth side of the rectangle does not require fence because of the barn.) We can solve this equation for h and substitute into the area equation:

$$
\begin{aligned}
x + 2h &= 100 \\
\Rightarrow 2h &= 100 - x \\
\Rightarrow h &= 50 - \frac{x}{2} \\
A(x) &= xh \\
&= x\left(50 - \frac{x}{2}\right) \\
&= 50x - \frac{x^2}{2}
\end{aligned}
$$

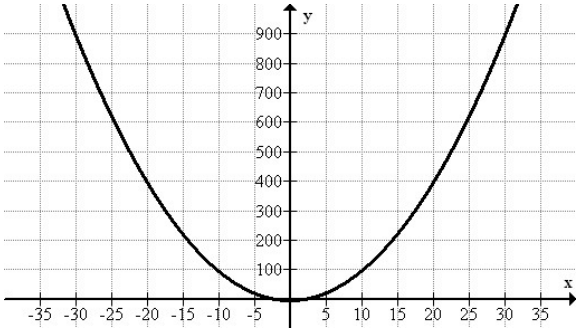
The graph of $A(x)$ is shown here on the interval [0,100]. Using a maximum function on a graphing utility tells us that the point (50,1250) is the maximum point. This tells us that when the rectangle's width is 50 ft, the area is 1250 ft^2.



In both of the previous two examples, we have considered a function on a subset of its domain. In particular, we limited our graphs to positive x values, as we were dealing with real quantities that could not reasonably be negative. However, if you examine the functions $S(x) = \frac{48}{x} + 2x^2$ and $A(x) = 50x - \frac{x^2}{2}$ for all real values of x, you will notice a difference in their graphs: the point we identified on $A(x)$ is the highest point the function has on its entire domain. This is not the case for $S(x)$. In the next lesson we will consider more functions like $S(x)$. In the remainder of this lesson we will consider functions like $A(x)$.

Symbolic definition of minimum

Consider the function $f(x) = x^2 - 5$. If you graph this function, you should see that the point (0, -5) is the lowest point on the graph. Even if we greatly increase the interval on which we view the graph, this point will always be the lowest point.



This point is referred to as the minimum value of the function. This minimum value is also referred to as a **global minimum,** to emphasize the fact that it is the lowest point on the entire graph.

We formally define the minimum point of a function as follows: The point $(c, f(c))$ is the minimum value of a function if $f(c)$ $f(a)$ for all elements a (a c) of the domain of f. Simply put: the point $(c, f(c))$ is the minimum if all other function values are greater than or equal to $f(c)$.

We can define the maximum of a function in a similar manner.

Symbolic definition of maximum

The function $A(x) = 50x - \frac{x^2}{2}$ has a **global maximum** at the point (50, 1250). Formally we can define the maximum of a function as follows: The point $(c, f(c))$ is the maximum of $f(x)$ if $f(c)$ $f(a)$ for all elements a (a c) of the domain of f. Simply put: the point $(c, f(c))$ is the maximum if all other function values are less than or equal to $f(c)$. Together, maximum and minimum points are called **extrema.** It is important to note that not all functions have extrema.

Example 4: Determine if each function has a minimum or a maximum point

Table 2.10:

a. $y = 2x$ - 1	b. $y = x^4$

Solution

a. The graph of $y = 2x$ - 1 is a line. It does not have a maximum or a minimum.

b. The graph of $y = x^4$ has a minimum value at (0,0). It does not have a maximum.

Lesson Summary

In this lesson we have focused on the idea of extrema: the minimum and maximum values of a function. Throughout the lesson we have examined examples of situations that involved quantities to be maximized or minimized. For example, we found the minimum surface area of a box, given a fixed volume, and we found the maximum area of a rectangular enclose, given a fixed perimeter. In order to find extrema, it is necessary to consider the graphs of functions. It is important to note again that when you study calculus, you will learn algebraic methods for finding exact values of extrema. For now we have used graphs to approximate these values. It is also important to note that not all functions have global extrema. In the next lesson we will consider functions (like the surface area example in this lesson) that have extrema that are not "global".

Points to Consider

1. What information might you look for in a situation in order to determine if the quantity involved should be maximized or minimized?
2. What do you think are the pros and cons of using a graphing utility to identify extrema?
3. What aspects of the equations of functions might indicate that a function has a maximum or a minimum value?

Review Questions

1. What quantity should be maximized? What quantity should be minimized? You are manufacturing chairs, and it costs you a certain amount of money to make each chair.
   You need to determine the selling price of the chairs.
2. A rectangle has area 20 in^2. Write an expression for the perimeter of the rectangle as a function of its width x.
3. What dimensions of the rectangle in problem #2 will minimize its perimeter? What is the minimum perimeter? (These values will be approximations.)
4. In your own words, define the term "maximum of a function."
5. Explain how you can use a graph to identify global extrema of a function.
6. A rectangle has a perimeter of 24 inches. What is the maximum area the rectangle can have?
7. A cylindrical canister has a volume of 30 in^3. What is the radius of the canister with minimum surface area? (Volume of a cylinder is $V = \pi r^2 h$
8. Consider the function $f(x) = bx^2 + 7$. For what values of b will the function have a maximum?
9. Consider the function $S(x) = \frac{48}{x} + 2x^2$. How can you tell that this function does not have a global maximum or minimum?
10. A rectangle has perimeter P. Write a function for the area of the rectangle as a function of P and x, the width of the rectangle. What do you think will be the rectangle with maximum area?

Review Answers

1. The cost per chair should be minimized. The profit (a function of the selling price) should be maximized.
2. $P = 2x + \frac{40}{x}$
3. When x=4.472, the perimeter is about 17.889 feet.
4. Answers will vary. Example: A maximum is the highest point on the graph, or the greatest y value of the function.
5. Answers will vary. Example: After graphing a function, you need to look for the highest or lowest point on the graph. It is important to explore a large subset of the domain, unless you are very familiar with a particular kind of function.
6. The rectangle with maximum area is a 6x6 square, with area 36 in^2.
7. $r \approx 1.68$ inches.
8. If b<0, the function will have a maximum.
9. If you examine the graph for negative x values, you see that the graph has more than one "piece". The piece that we did not consider in the example in the text goes above and below the point we identified in the example in the text.
10. $A(x) = x\left(\frac{P}{2} - x\right)$If you choose values of P, and examine the graphs, you will see that the rectangle with maximum area is always a square, like the one in problem 6.

2.3 Increasing and decreasing functions

Learning objectives

- State intervals on which a function is increasing or decreasing.
- Do this in the context of discrete and non-discrete functions.
- Identify relative maxima and minima of a function by considering the graph of a function as well as function values represented in a table.

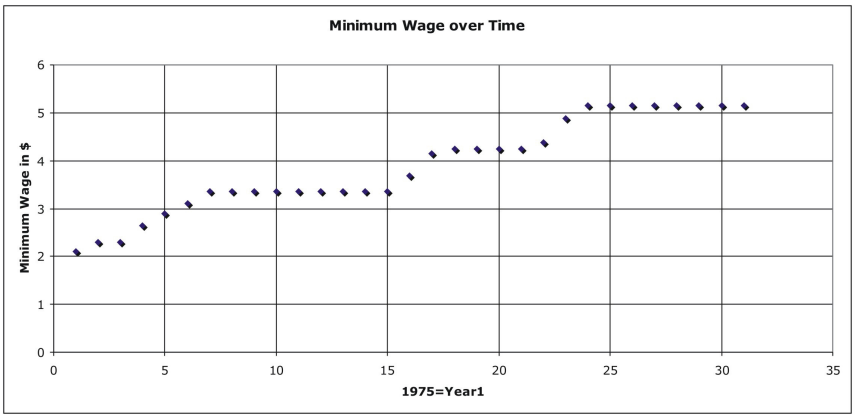
Introduction

In the previous lesson we examined maxima and minima, the highest and lowest values of a function. In this lesson we will consider the function values in between the extrema, focusing on where function values are **increasing** or **decreasing.** A function is increasing on some interval of its domain if f(a) > f(b) for all a, b in that interval such that $a > b$. A function is decreasing on some interval of its domain if $f(a) < f(b)$ for all a, b in that interval such that $a > b$. More casually put: a function is increasing if the graph rises to the right. It is decreasing if the graph falls to the right.

We will consider these ideas in the context of two different kinds of functions: discrete and non-discrete functions, and we will also return to the type of function we saw in the previous lesson that had a minimum value that was not a global minimum.

Discrete Functions

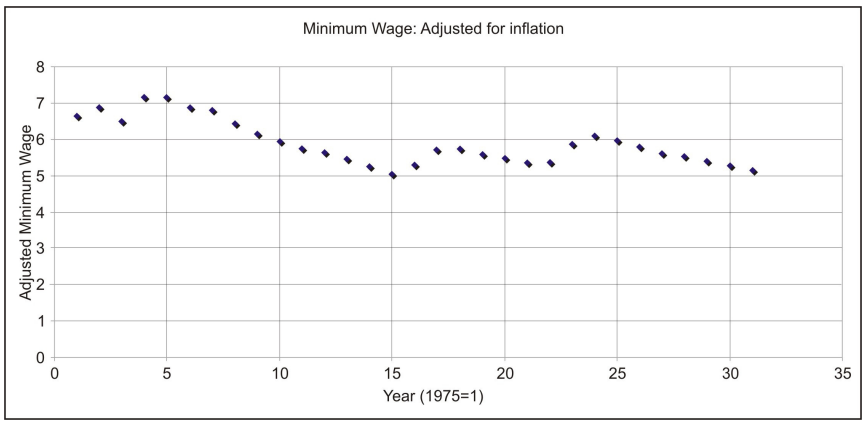
A discrete function is a function in which the domain and range are each a discrete set of values, rather than an interval in $\mathbb{R}$. Consider for example the graph below, which shows the federal minimum wage from 1975 to 2005:



Notice that overall, from 1975 to 2005, the minimum wage (in dollars per hour) increased from $2.10 per hour to $5.15 per hour. However, if we examine the graph on smaller intervals, we can see that this increase was not steady. For example, the minimum wage increased from $2.30 in 1977 (year 3) to $3.35 in 1981 (year 7), and then it was constant for several years at $3.35 per hour. We can describe this situation using intervals. For example, we can say that the function increases on the interval (3,7) and is constant on the interval (7, 15).

This pattern of increasing for several years and then being constant for several years continues until the end of the data set in 2005. (The minimum wage increased to $5.85 in 2007.) Notice, however, that the function never decreases. In example 1 you will consider a function that decreases.

Example 1: The graph below shows the minimum wage data *adjusted for inflation.* (Since prices for food and other necessities tend to go up over time, comparing absolute dollar amounts doesn't make sense. The adjustment shows how much one can buy with minimum wage.) State the intervals on which the graph increases and the intervals on which it decreases.



Solution:

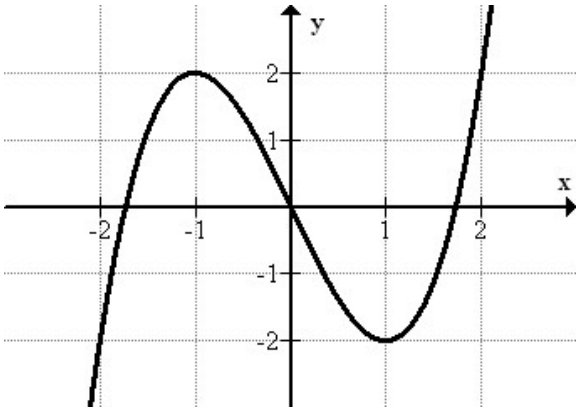
The function decreases on the intervals (2, 3), (5, 15), (18, 21), (24, 31).
The function increases on the intervals (1, 2), (3, 4), (15, 17), (22, 24).

Notice that the intervals are ranges of specific points. That is, the function is decreasing if, when we look at the graph, the points are sloping down from left to right. The function is increasing if the points are sloping up from left to right. This will be the case for any discrete function. Next we will consider what it means for a non-discrete function to be increasing or decreasing.

Non-Discrete Functions

A non-discrete function is one that is **continuous** either on its entire domain, or on intervals within its domain. The term continuous refers to a function whose graph has no holes or breaks. (Note that this is not a formal definition. To formally define continuity requires that we use the concept of limit, which we will examine in the next lesson. For now it is sufficient to focus on what the graph looks like.)

The graph below shows a function, $y = x^3 - 3x$, that is continuous on its entire domain, which is $\mathbb{R}$. The function increases and decreases on different intervals within its domain.



The function increases on the interval $(-\infty, -1)$ and on the interval $(1, \infty)$. The function decreases on the interval $(-1, 1)$. (The reason these are open intervals (with parentheses instead of brackets) is because the function is neither increasing nor decreasing at the moment it changes direction. We can imagine a ball thrown into the air. The height increases up to a maximum point before it starts to decrease. What happens at that maximum point? There is an instant when the height is not changing; it is neither increasing nor decreasing. This is why the maxima and minima are not included in these intervals.)

Not every function will increase *and* decrease. For example, the linear function $y = 2x + 1$ increases on its entire domain, while the linear function $y = -2x + 1$ decreases on its entire domain. In example 2, you will consider graphs and determine where the function increases and where it decreases.

Example 2: State the intervals on which each function increases and decreases.

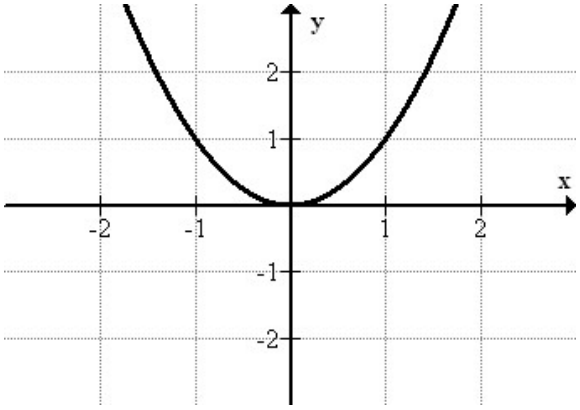
Table 2.11:

a. $y = x^2$	b. $y = x$

Solution:

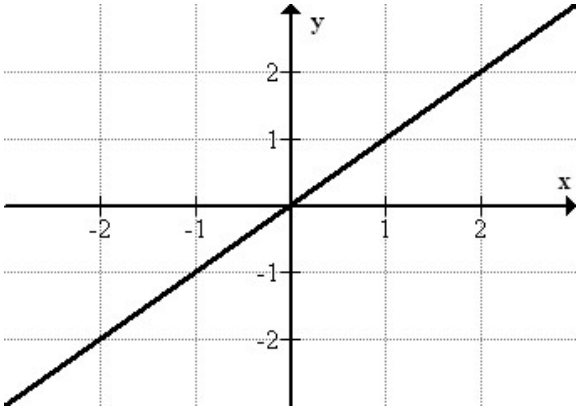
a. $y = x^2$

The graph of this function is shown here. The function decreases on $(-\infty, 0)$ and it increases on $(0, \infty)$.



b. $y = x$

The graph of this function is shown below. This function is a line with positive slope. That is, it rises to the right. Therefore the function is increasing on its entire domain, $\mathbb{R}$.



Notice that the function $y = x^2$ has a global minimum at (0, 0), and the function y = x has no extrema. However, the function $y = x^3 - 3x$, much like the surface area equation in the previous lesson, has a maximum and minimum value, if we only consider the function on certain intervals within its domain. Next we will consider these kinds of extrema.

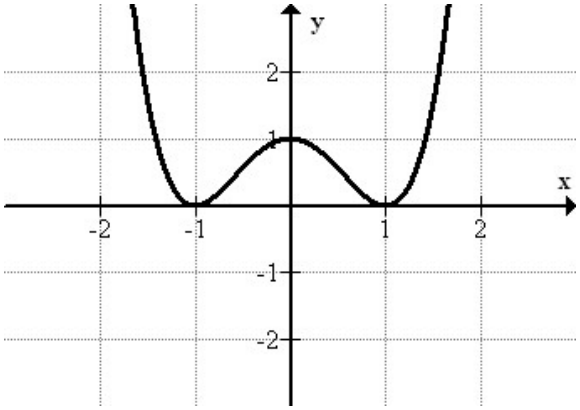
Relative Maxima and Minima

Consider again the function $y = x^3 - 3x$. This function has no global extrema. However, consider the function values on the interval (-2, 0). On this interval the point (-1, 2) is a maximum. We refer to this kind of maximum as a **relative maximum** or a **local maximum**. Informally, you can think of a relative maximum as the highest point on a graph in some interval of the domain. Formally, a function value $f(c)$ is a relative maximum of a function $f(x)$ if $f(c) > f(a)$ for all a in some interval of the domain of $f(x)$. Note that for $y = x^3 - 3x$, we identified the interval (-2, 0) around the point (-1. 2). However, this interval is not unique. We could have identified a larger or smaller interval around (-1, 2). What matters is that this interval exists.

Now consider the function $y = x^3 - 3x$ on the interval (0, 2). The point (1, -2) is a minimum on this interval. We refer to this kind of minimum as a **relative minimum** or a **local minimum**. Informally, you can think of a relative minimum as the lowest point on a graph in some interval of the domain. Formally a function value $f(c)$ is a relative minimum of a function $f(x)$ if $f(c) < f(a)$ for all a in some interval of the domain of $f(x)$. Again, while we identified a particular interval in which the point (1, -2) is a minimum, what matters is that such an interval exists.

Notice that the function values before a maximum are increasing, and after they are decreasing. The opposite is true for a minimum. In the following example, you will identify both global and relative maxima and minima.

Example 3. Identify any global maxima and minima, as well as relative maxima and minima, for the function shown below.



Solution:

This function has no global maxima, as both ends continue to go up infinitely. The points (-1, 0) and (1, 0) together are the global minima of the function. The point (0, 1) is a relative maximum.

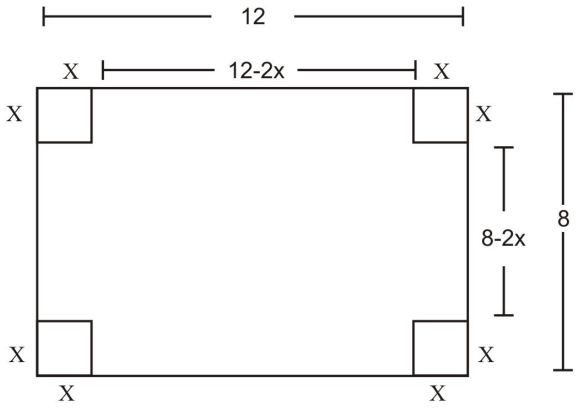
Now that we have defined the terms relative maximum and relative minimum, we can justify the decision in the previous lesson to examine the function $S(x) = 48/x + 2x^2$ on the interval [0, 5]. The function has no global extrema, but the point (2.3, 31.4) is an approximation of a relative minimum of the function. Next we will consider a similar problem, focusing on how we can make use of technology to help identify extrema.

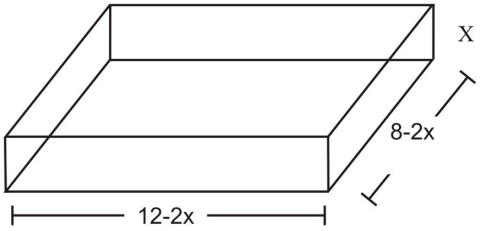
Applications and Technological tools

For all of the examples we have considered so far, we have looked at the graph of a function in order to identify a maximum or a minimum. The following example uses this technique:

You have a rectangular piece of cardboard that has dimensions 12 inches by 8 inches. You cut square pieces out of each corner so that you can fold the cardboard into an open box. What are the dimensions of the box with maximum volume?

First we need to represent the volume as a function. If x represents the side length of each square cut out of a corner, then the dimensions of the box will be 12 - $2x$, 8 - $2x$ and x. The pictures below show the flat cardboard, and the box:





The volume of the box is the product of its dimensions. Therefore we can write an equation that expresses the volume as a function of x: $V(x) = (12 - 2x)(8 - 2x)x$.

Now we need to look at the graph of this function. There are several options for creating a graph, but the most common tool is a TI graphing calculator. Enter the equation into the "$Y =$" menu. (Be sure to clear any equations already in $Y =$.) To view the graph you need to set an appropriate window. One way to determine a good window is to look at a few values in a table. To set the table, press 2^{nd}, WINDOW. You can either scroll through a table of values, or you can choose x values. To choose x values, set Indpnt (x is the "independent variable) to "ask." Now press 2^{nd}, GRAPH and enter x values into the table. The table below shows several values for x and $V(x)$:

Table 2.12:

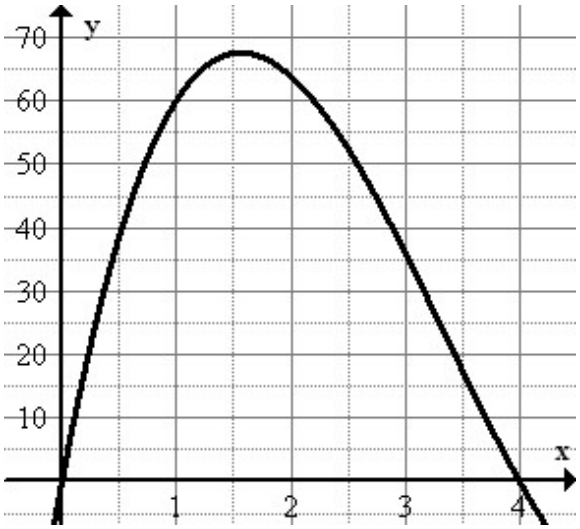
x	$y = V(x)$
0	4
1	60
2	64
3	36
4	0
5	
6	0
7	84

Table 2.12: (continued)

x	$y = V(x)$
8	256

The values in the table might convince you that the function values increase without bound, and hence that there is no maximum. However, we have to keep in mind the context of the problem. We can only cut so much out of each side of the cardboard to make the squares. For the side with length 8, if we cut out two squares of side length 4, there will be no cardboard left to fold into a box. Therefore we should consider the function on the interval (0, 4).

To set the viewing window, press WINDOW. Set Xmin to 0 and Xmax to 4. Set Ymin to 0 and Ymax to some value larger than 64. (The table above tells us that the function values go up to at least 64.) You may also want to set the Y-scale to 10 so that the graph does not show too many tick marks. Now press GRAPH to view the graph of the function.



The graph should show a clear maximum between $x = 1$ and $x = 2$. To find an approximation of this point, press 2^{nd}, TRACE, which takes you to the CALC (for "calculation") menu. Choose option 4, Maximum. This brings you back to the graph. You will be asked to input a left-bound for the value. Scroll to the left of the maximum, or enter a value for x that is clearly to the left of the maximum, and then press enter. Then do the same thing for a right-bound. When you see the word "Guess", press Enter. The calculator will then take a moment to approximate the coordinates of the maximum point.

Your calculator should tell you that the point is approximately (1.57, 67.60). This means that if you cut out squares with side length approximately 1.57 in, the box will have a maximum volume of about 67.60 in^3.

Note: Other graphing utilities may have this function built in, but the procedure may be slightly different.

Table 2.13:

Technology Note
The TI-83/84 calculators can be used to create a table of selected values as follows

- Enter the function by pressing <TI font_Y=> (see below)
- Set the table input to ***Ask***, as follows:

◦ Press <TI font_2nd>[TBLSET]and choose ***Ask*** for the independent variable, as shown below
◦ Press <TI font_2nd>[TABLE], and enter values as shown below

Note: Any spreadsheet program is good for creating tables like this as well.

Lesson Summary

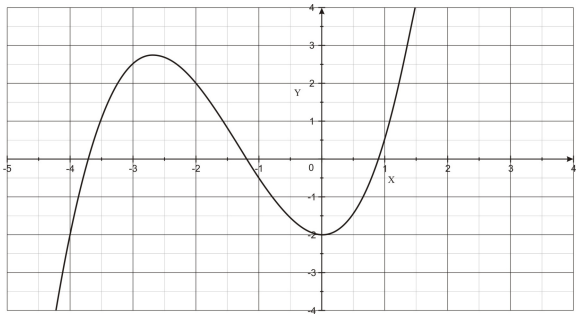
In this lesson we have examined the behavior of functions between its maxima and minima. We have formally and informally defined what it means for a function to increase or decrease, and we have identified intervals on which specific functions increase or decrease. We have also defined the terms relative, or local maximum and minimum. The idea of local maxima and minima allow us to solve max-min problems by considering a function on a particular, relevant sub-interval of its domain, rather than on its whole domain. This is necessary, given the reality of situations. For example, many of the quantities in models of real life problems cannot be negative. If we consider the graph of a function on a relevant interval, we can approximate relative maxima and minima by inspecting a graph, or by using the functions of a graphing utility, such as a max-min function, or table of values.

Points to Consider

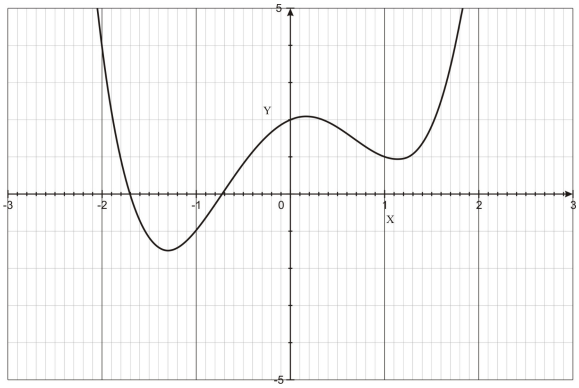
1. What is the difference between a discrete and non-discrete function?
2. What would the graph of a function look like if the function is always increasing, or always decreasing?
3. Do all functions increase or decrease?

Review Questions

1. Give an example of a discrete function that increases and decreases on different intervals of its domain.
2. For questions 2 and 3, consider the following graph:

3. 

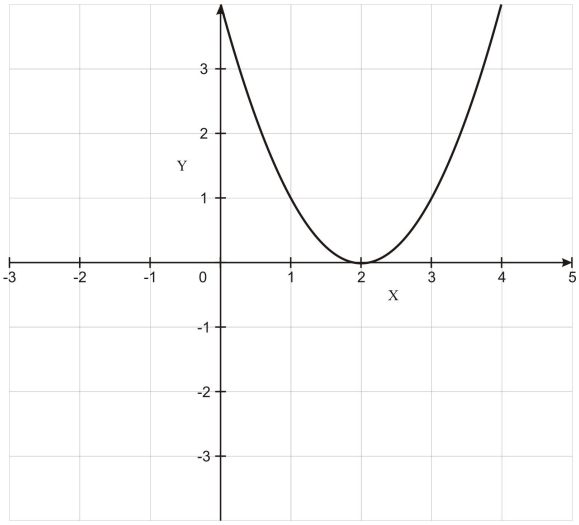
4. Approximate the coordinates of the relative maximum of the graph
5. Approximate the coordinates of the relative minimum of the graph.
6. Consider the equation of a linear function $y = mx$, where m is the slope of the line. For what values of m will the function be an increasing function? For what values of m will the function be a decreasing function?
7. Consider the graph above for questions 2 and 3. On what intervals is the function increasing? On what intervals is the function decreasing?
8. Sketch a possible graph of the function described here: The function is continuous on $\mathbb{R}$. It is decreasing on the interval $(-\infty, 2)$ and increasing on the interval $(2, \infty)$.
9. Explain in your own words the difference between relative and global extrema.
10. For the function shown in the graph below, give approximate coordinates of all global and relative extrema.



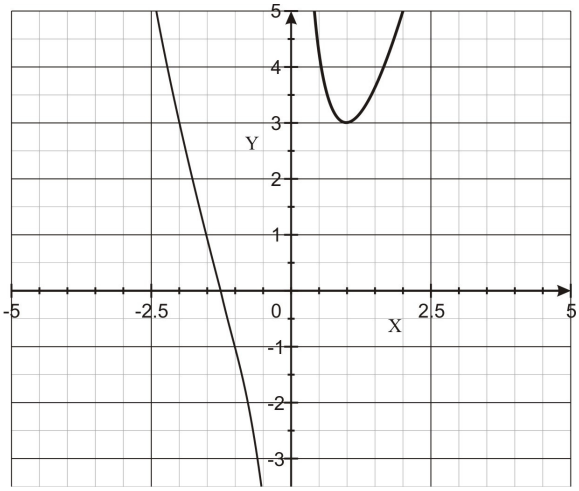
11. Consider the function $f(x) = x^2 + \frac{2}{x}$. Use a graphing utility to sketch a graph of this function and to calculate a relative minimum.
12. A rectangle is inscribed in a semi-circle of radius 3. a. Write an equation to represent the area of the rectangle as a function of x.
    b. Graph the equation and calculate the maximum area of the rectangle.

Review Answers

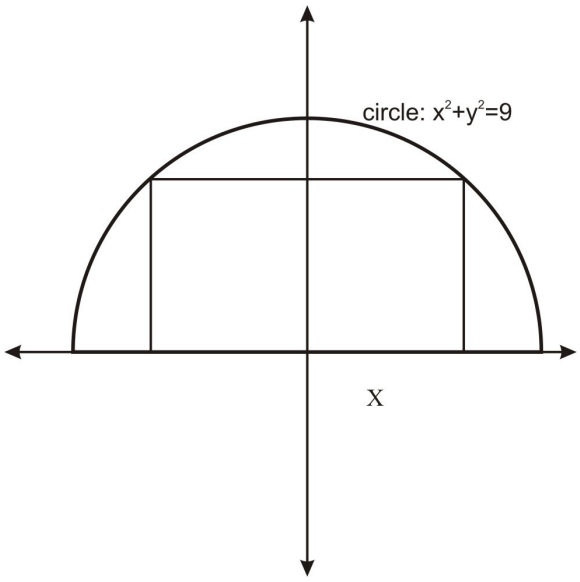
1. Answers will vary. Example: a person's height as a function of age. The function values will increase until the age at which the person stops growing. Then the function will be constant for several decades. Then the function will decrease slightly, as some people lose height in their later years.
2. (-2.7, 2.75)
3. (0, -2).
4. If $m > 0$, the function will be increasing.If m < 0, the function will be decreasing
5. Increasing: $(-\infty, -27)$, $(0, \infty)$. Decreasing: (-2.7 , 0).
6. Answers will vary. Example:

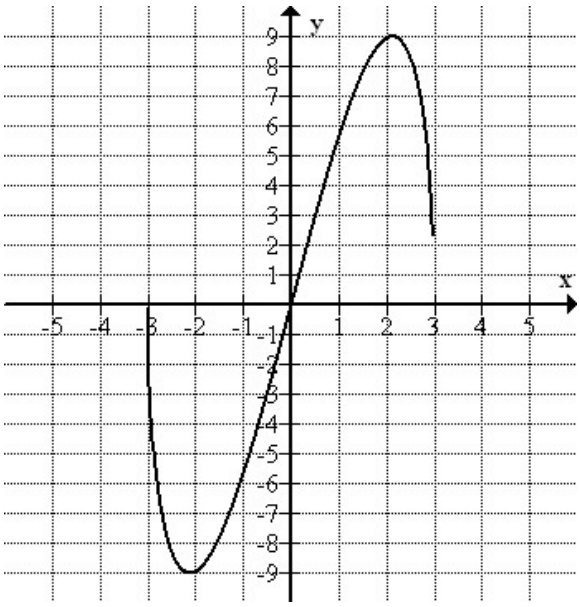


7. Answers will vary. Example: a global maximum is the greatest function value, or the highest point on the graph of the function. A global minimum is the least function value, or the lowest point on the graph of a function. A relative max or min is the highest or lowest point within some interval in the domain of the function.
8. The function has a global minimum at approximately (-1.3, -1.5), a relative maximum at approximately (0.17, 2.0), and a relative minimum at (1.1,1).
9. The graph has a relative minimum at (1, 3).



10. a. $A = 2x\sqrt{9 - x^2}$ b.





The graph has a maximum at (2.12, 9). So the maximum area of the rectangle is 9.

2.4 End Behavior of Functions

Learning objectives

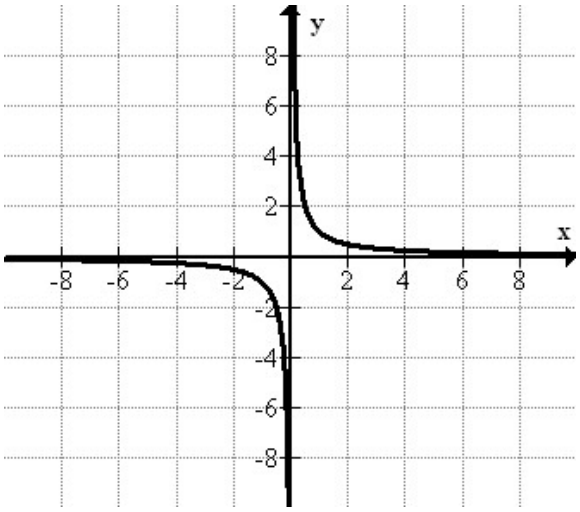
- Determine the end behavior of different kinds of functions, including those that approach a particular value, those that approach infinity, and those that follow neither pattern.
- Use the concept of limit to describe the end behavior of functions.
- Use and interpret limit notation to describe the end behavior of functions.
- Identify simpler functions that can be used to describe the end behavior of more complicated functions.

Introduction

In the previous lesson we considered the behavior of functions on intervals within the functions' domains. In particular, we identified intervals on which the functions increased or decreased, and we identified points that were relative maxima or minima. Now we will analyze the **end behavior** of functions, which refers to the behavior of functions as we let x be large positive or negative values. Essentially, we will be considering the "ends" of the x-axis, the far right and left of the graph.

End behavior and the concept of limit

Consider the function $f(x) = \frac{1}{x}$. A graph of this function is shown below.



Notice that as the values of x get larger and larger, the graph gets closer and closer to the x-axis. In terms of the function values, we can say that as x gets larger and larger, $f(x)$ gets closer and closer to 0. Formally, this kind of behavior of a function is called a **limit.** We say that as x approaches infinity, the limit of the function is 0.

If we then consider the behavior of the function as x approaches $-\infty$, we see the same behavior: the limit of the function has x approaches $-\infty$ is also 0.

To be even more formal, we can write these limits using a special notation. For the first limit, we write: $\lim_{x\to\infty} f(x) = 0$. For the second limit, we write $\lim_{x\to-\infty} f(x) = 0$. For any limit, here we will always write the x under the abbreviation "lim", and then we will write the function under consideration. We can also write each of these limits with the specific function:

$\lim_{x\to\infty} \frac{1}{x} = 0$ and $\lim_{x\to-\infty} \frac{1}{x} = 0$.

Example 1: Write the limit described using limit notation.

The limit of some function $f(x)$ as x approaches infinity is 2.

Solution: We write the limit as follows:

$\lim_{x\to\infty} f(x) = 2$

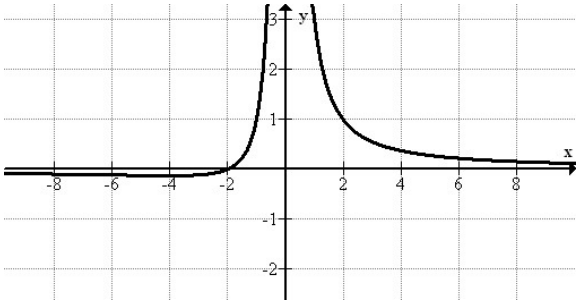
The function $f(x) = \frac{1}{x}$ shows just one of several types of end behavior. Next we will examine this type of end behavior in detail, and then we will consider two other kinds of end behavior.

Limit, asymptote

For the function $f(x) = \frac{1}{x}$, if x approaches ∞ or if x approaches $-\infty$, the graph of the function gets closer and closer to the x-axis, which is the line $y = 0$. The line $y = 0$ is called the **asymptote** of the graph. We can also say that $f(x) = \frac{1}{x}$ is **asymptotic** to the line y=0.

It is important to note two key points about asymptotes. First, because we are focused on end behavior, we are considering the limit of functions as x approaches $\pm\infty$, and so the asymptotes we will find are horizontal lines. If we were examining other aspects of functions, we might find asymptotes that are vertical lines. For example, the function $f(x) = \frac{1}{x}$ has a vertical asymptote at $x = 0$, or the y-axis. That is, the graph approaches the y-axis, as x values get closer and closer to 0.

The second point is about the relationship between an asymptote and the values of a function. For many functions, there is no function value equal to the asymptote. For example, the function $f(x) = \frac{1}{x}$ never takes on the function value 0. However, the graphs of some functions do cross an asymptote. Consider for example the function $f(x) = \frac{x+2}{x^2}$, shown in the graph below.



The graph of the function is asymptotic to the line $y = 0$ (i.e., the x-axis.) However, at x = -2, the function value is 0. That is, the function crosses the asymptote. The line is still an asymptote because as x approaches $-\infty$, the function values get closer and closer to 0.

Example 2. Determine the horizontal asymptote of the function $g(x) = \frac{2x-1}{x}$ and express the asymptotic relationship using limit notation.

Solution:

This function is asymptotic to the line $y = 2$. The limit is written as $\lim_{x\to\neq\infty} \frac{2x-1}{x} = 2$.

We can determine the asymptote (and hence the limit) if we look at the graph. However, we can also analyze the equation to determine the limit. Consider the function $g(x) = \frac{2x-1}{x}$. As x approaches infinity, the x values are getting larger and larger. For sufficiently large values of x, the values of the expression $2x$ - 1 are very close to the values of the expression $2x$, because subtracting one from a large number is fairly insignificant. Thus for sufficiently large values of x, $\frac{2x-1}{x} \approx \frac{2x}{x} \approx 2$. As you can see from the accompanying table, which was created by a TI-83 graphing calculator, the function value gets closer to 2 as we look at larger and larger x values.

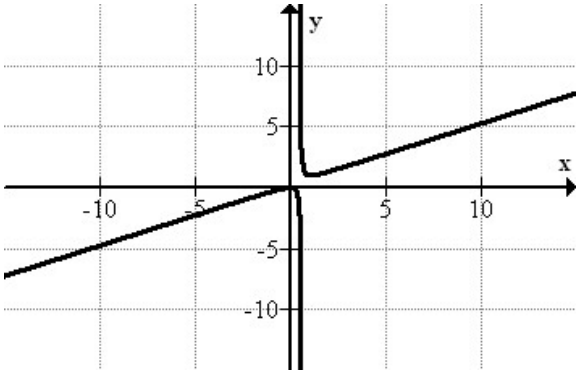
X	Y1
0	ERROR
50	1.98
100	1.99
150	1.9933
200	1.995
250	1.996
300	1.9967

X=0

While the functions we have considered so far both have a horizontal asymptote, not every function has a horizontal asymptote. Next we will examine such functions.

Infinite limits

Above we looked at the behavior of the function $g(x) = \frac{2x-1}{x}$. Now let's consider another function, $h(x) = \frac{x^2}{2x-1}$. The equation looks similar, but as you can see from the graph, the end behavior is quite different.



As x approaches ∞, the function values also approach ∞. Therefore the limit of the function as x approaches ∞ is: $\lim_{x\to\infty} \frac{x^2}{2x-1} = \infty$. Similarly, as x approaches $-\infty$, $f(x)$ approaches $-\infty$. Therefore we have $\lim_{x\to-\infty} \frac{x^2}{2x-1} = -\infty$.

We can also understand this limit if we analyze the equation for $h(x)$. As x gets larger and larger, the value of the expression $2x$ - 1 get closer and closer to the value of the expression $2x$. That is, for sufficiently large values of x, $2x$ - 1 $\approx$ $2x$. Therefore the values of $h(x)$ approach $\frac{x^2}{2x} = \frac{x}{2}$. As x gets larger and larger, so does $\frac{x}{2}$. The accompanying table shows that for large values of x, the function $h(x)$ gets closer and closer to $\frac{x}{2}$. Therefore the limit is infinity. We could make a similar argument for $\lim_{x\to-\infty} \frac{x^2}{2x-1} = -\infty$.

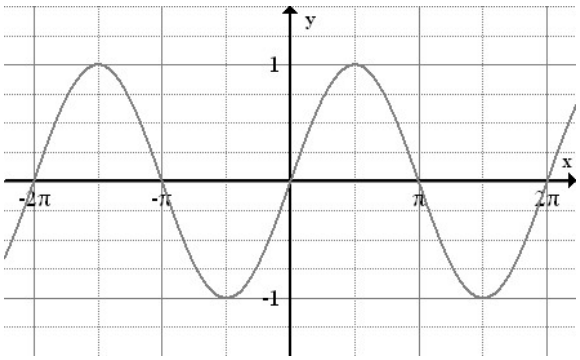
X	Y1
0	0
50	25.253
100	50.251
150	75.251
200	100.25
250	125.25
300	150.25

X=0

The functions $g(x) = \frac{2x-1}{x}$ and $h(x) = \frac{x^2}{2x-1}$ are both members of a family of functions known as rational functions. In chapter 2 you will study this family of functions, including their asymptotes. For now what is important for you to note is that some rational functions will have a horizontal asymptote, while others will not. A function like $g(x)$ will always have a horizontal asymptote that you can determine by analyzing the graph or the equation, as we have done above. A function like $h(x)$ will not have a horizontal asymptote, but instead will grow without bound. That is, the limit will be $+\infty$ or $-\infty$. Other functions may show neither pattern.

Non-existent limit

The functions that we have examined so far have shown one of two patterns: as x approaches $\pm\infty$, the functions approach some limit L, or the function approaches $+\infty$ or $-\infty$. The function shown below shows neither pattern.



As x approaches ∞, the function values oscillate between -1 and +1, but they never approach 0, or any other number. As x approaches $-\infty$, the same behavior occurs. Therefore $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ do not exist.

Example 3: Describe the end behavior of each graph. That is, determine if the function has a limit L, if the limit is infinite, or if the limit does not exist.

a. $y = x^2$
b. $y = 2(-1)^x$
c. $y = 1 - \frac{1}{|x|}$

Solution:

a. $y = x^2$

As x approaches $+\infty$, x^2 also approaches $+\infty$. As x approaches $-\infty$, x^2 approaches $+\infty$. Therefore $\lim_{x\to\infty} x^2 = \lim_{x\to-\infty} x^2 = \infty$.

b. $y = 2(-1)^x$

This function is difficult to understand without producing a graph. The table shows that the function only takes on two values: 2, and -2, and is undefined at non-integer values of x. As x approaches $+\infty$ or $-\infty$, the function values alternate between 2 and -2. Therefore the limit does not exist.

X	Y1
0	2
.5	ERROR
1	-2
1.5	ERROR
2	2
2.5	ERROR
3	-2

X=0

c. $y = 1 - \frac{1}{|x|}$

If you look at the table of this function, which has negative and positive values of x, you can see that as x approaches $+\infty$ or $-\infty$, the function values approach 1.

X	Y1
-10000	.9999
-200	.995
-10	.9
0	ERROR
10	.9
200	.995
10000	.9999

X= -10000

Therefore $\lim_{x\to\infty}\left(1 - \frac{1}{|x|}\right) = \lim_{x\to-\infty}\left(1 - \frac{1}{|x|}\right) = 1$.

We can also determine this limit analytically. For large values of x, $|x|$ is also large, and so $\frac{1}{|x|}$ is small (since dividing 1 by a large number results in a very small number). Therefore, for large values of x, $1 - \frac{1}{|x|} \approx 1 - 0 = 1$. We can make the same argument for x approaching $-\infty$.

Thus far we have categorized functions in two ways: those that approach a horizontal line, and those that

do not. Of those that do not, some have an infinite limit, while some do not. Next we will consider the end behavior of functions that have an infinite limit.

Simpler functions

Consider again the function $h(x) = \frac{x^2}{2x-1}$. As we found above, as x approaches infinity, $h(x)$ also approaches infinity. Analyzing the equation, we also found that for large values of x, $h(x) \approx \frac{x}{2}$. The table below shows values of each function for larger and large values of x:

Table 2.14:

x	$h(x)$	$\frac{x}{2}$
1	1	0.5
10	5.26	5
100	50.25	50
1000	500.25	500

Given this relationship between h(x) and the line $y = \frac{x}{2}$, we can use the line to describe the end behavior of $h(x)$. That is, as x approaches infinity, the values of $h(x)$ approach $y = \frac{x}{2}$. As you will learn in chapter 2, this kind of line is called an **oblique asymptote,** or **slant asymptote**.

Some functions, however, may approach a function that is not a line. Consider for example the function $f(x) = x^2 + 2x - 3$. This function has an infinite limit as x approaches infinity. However, as x gets larger and larger, $f(x) \approx x^2$. Therefore we can use the function $y = x^2$ to describe the end behavior of $f(x)$.

Lesson Summary

In this lesson we have focused on the end behavior of functions. End behavior refers to the behavior of the function as x approaches $-\infty$ or as x approaches $-\infty$. In terms of the graph of a function, analyzing end behavior means describing what the graph looks like as x gets very large or very small. Some functions approach certain limits.

Graphically, this means the function has a horizontal asymptote. As x approaches ∞ or $-\infty$, the function gets closer and closer to a horizontal line. In this lesson we identified horizontal asymptotes graphically and algebraically. Other functions have infinite limits. For these functions, as x approaches ∞ or $-\infty$, the function also approaches ∞ or $-\infty$. That is, the graph increases or decreases without bound. Finally, some functions have no limit as x approaches ∞ or $-\infty$.

It is important to note that you will return to the concept of limit again in your study of functions, especially when you study calculus. Here we have only focused on one kind of limit, that of x approaching ∞ or $-\infty$. In your later studies you will also examine the limit of a function as x approaches a particular finite value.

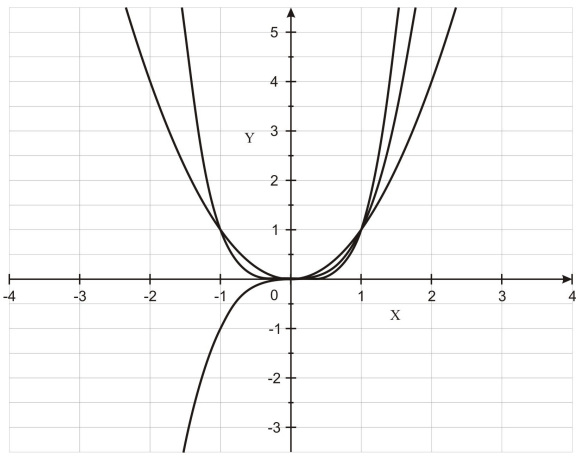
Points to Consider

1. What does it mean for a limit to exist, or for a limit to not exist?
2. How do graphs help us analyze end behavior and limits?
3. What kinds of graphs have horizontal asymptotes?
4. What kinds of graphs might have a vertical asymptote?

Review Questions

1. Explain in words the meaning of the limit statement: $\lim_{x\to\infty}\left(3+\frac{2}{x}\right)=3$
2. Write a limit statement for the following description of the end behavior of a function: As x approaches negative infinity, the values of the function $f(x)=\frac{1}{x}$ approach 0.
3. Describe the end behavior of the function $y=\frac{1}{x^2}$ using limits.
4. Describe the end behavior of the function $y=\frac{5x}{x-2}$ using limits.
5. Describe the end behavior of the function $y=\frac{ax}{x-2}$ using limits and explain the limits using an algebraic argument.
6. Describe the end behavior of the function $y=\frac{x^2}{x-5}$.
7. Describe the end behavior of the function $f(x)=\cos(x)$.
8. Explain in your own words the definition of a horizontal asymptote.
9. Consider the functions $y=x^2$, $y=x^3$, and $y=x^4$. The functions $y=x^2$ and $y=x^4$ are called even functions. Even functions are functions for which $f(x)=f(\text{-}x)$. The function $y=x^3$ is an odd function. Odd functions are functions for which $f(\text{-}x)=\text{-}\ f(x)$a. Graph these three functions using a graphing calculator and describe their end behavior.
   b. Describe the end behavior of odd and even power functions. (A power function is one of the form $y=x^n$ for integer values of n.)
   c. Now describe the end behavior of the function $y=x^{17}$.
10. Describe the end behavior of the function $f(x)=\frac{1}{2}x^2-4x-1$ using a simpler function.

Review Answers

1. As x approaches infinity, the value of the expression $3+\frac{2}{x}$ approaches 3.
2. $\lim_{x\to-\infty}\left(\frac{1}{x}\right)=0$
3. $\lim_{x\to-\infty}\left(\frac{1}{x^2}\right)=0, \lim_{x\to\infty}\left(\frac{1}{x^2}\right)=0$
4. $\lim_{x\to-\infty}\left(\frac{5x}{x-2}\right)=5, \lim_{x\to\infty}\left(\frac{5x}{x-2}\right)=5$
5. $\lim_{x\to-\infty}\left(\frac{ax}{x-2}\right)=a, \lim_{x\to\infty}\left(\frac{ax}{x-2}\right)=a$As the value of x gets very large, the value of the expression x - 2 is very close to x. So the expression $\frac{ax}{x-2}\approx\frac{ax}{x}=a$. The same is true as x approaches negative infinity.
6. As x approaches infinity, the function values approach infinity. As x approaches negative infinity, the function values approach negative infinity
   The graph is also asymptotic to the function $y=x$.
7. The function has no limit as x approaches infinity or negative infinity.
8. A horizontal asymptote is a line that a function approaches as x approaches infinity or negative infinity. Not every function has one.
9. 

The graphs of $y=x^2$ and $y=x^4$ have the same end

behavior: both ends approach infinity. The graph of the function $y = x^3$ is different. As x approaches infinity, the function values approach infinity. As x approaches negative infinity, the function values approach negative infinity. (One end goes up, one goes down!)
b. If the exponent is even, the end behavior is that of $y = x^2$ and $y = x^4$: both ends go up. If the exponent is odd, the end behavior is that of $y = x^3$: one end approaches infinity, and the other end approaches negative infinity.
c. One end will approach infinity, and the other end will approach negative infinity.

10. As x gets large, $f(x) \approx \frac{1}{2}x^2$As x approaches $\pm\infty$, both functions approach ∞.

2.5 Function families

Learning objectives

- Determine if an equation represents a particular family of functions.
- Determine if a graph represents a particular family of functions.
- Identify key aspects of different families of functions.
- Sketch basic graphs of families of functions.

Introduction

In the previous lessons we have worked with many different functions and their graphs. From the examples in the lessons, as well as from your previous experience working with equations and graphs, you may have already made connections between the forms of equations of functions, and what the graphs look like. In this lesson we will examine several "families" of functions. A family of functions is a set of functions whose equations have a similar form. The "parent" of the family is the equation in the family with the simplest form. For example, $y = x^2$ is a parent to other functions, such as $y = 2x^2$ - $5x$ + 3.

Here we will focus on the key aspects of each family. In the next lesson we will focus on relationships within each family, and how we can use certain relationships to make graphing by hand more efficient.

Linear functions

An equation is linear if it contains no powers of x besides $x^1 = x$ or $x^0 = 1$. For example, $y = 2x$ and $y = 2$ are linear equations, while $y = x^2$ and $y = 1/x$ are non-linear.

Linear equations are called linear because their graphs form straight lines. As you may recall from your earlier studies of algebra, we can describe any line by its average rate of change, or slope, and its y-intercept. (In fact, it is the constant slope of a line that makes it a line!) These aspects of a line are easiest to identify if the equation of the line is written in slope-intercept form, or $y = mx + b$. The slope of the line is equal to the coefficient m, and the y-intercept of the line is the point (0, b).

Example 1: Identify the slope and the y-intercept of each line:

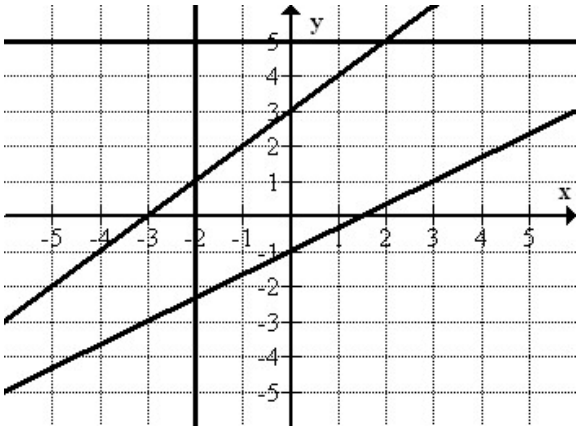
Table 2.15:

a. $y = (2/3)\ x$ - 1	b. $y = 5$	c. $x = -2$	d. $2x$ - $3y = 9$

Solution:

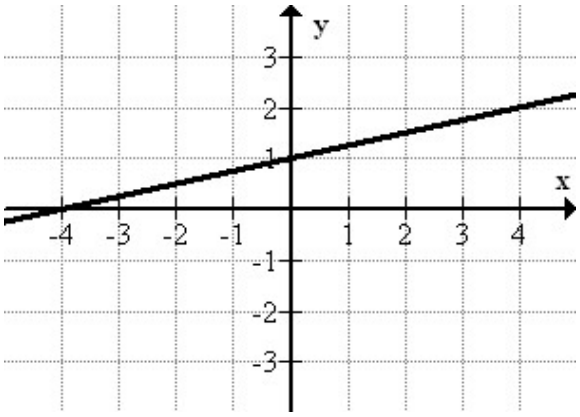
a. $y = (2/3)\ x$ - 1
This line has slope (2/3) and the y-intercept is the point (0, -1).
b. $y = 5$
This is a horizontal line. The slope is 0, and the y-intercept is (0,5).
c. $x = -2$
This is a vertical line. The slope is undefined, and the line does not cross the y-axis. (Note that this line is not a function!)
d. $y = 2/3x + 3$
The slope of this line is 2/3, and the y-intercept is the point (0, 3).

The graph below shows these four lines together:



Notice that the function value of $y = 5$ is 5 for all x values. This kind of function is referred to as a **constant function**, as the function values are constant, or unchanging. Together, the constant functions make up a "sub-family" of the linear functions. That is, they are all functions whose graphs are horizontal lines.

For the linear functions $y = (2/3)x$ - 1 and $y = x + 3$, it should be straightforward to identify slope and y-intercept. If the equation of a line is in another form, these aspects of the line are not necessarily obvious. Consider for example the linear equation $4y$ - $x = 4$. In this form, the coefficient on x is -1, but as the graph below shows, the slope is *not* -1, but 1/4.



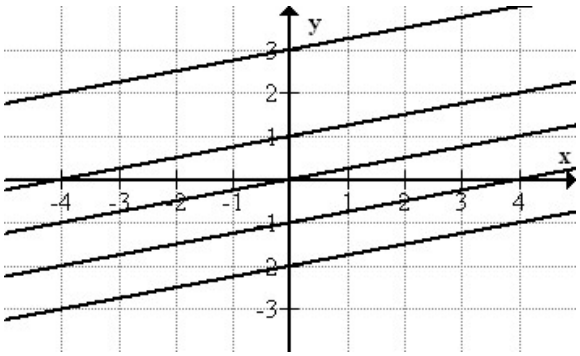
To identify the slope of the line, you can either rewrite the equation so that it is in $y = mx + b$ form, or you can examine the graph. For example, note that the graph of $4y$ - $x = 4$ passes through the points (0, 1) and (4, 2). We can see from these points that the slope is (1/4). If we rewrite the equation, we can verify the slope and the y-intercept:

$4y$ - $x = 4$

$4y = x + 4$

$y = (1/4)x + 1$

While we can think of this line as a member of the family of linear functions, we can also think of it as a member of a sub-family of linear functions with slope = 1/4. The graphs of this subfamily will be a set of parallel lines:

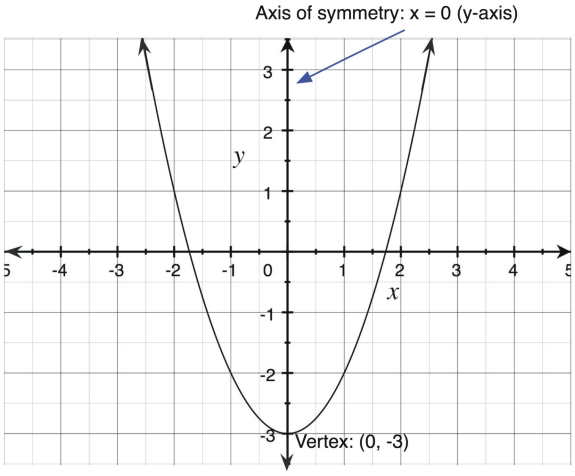


Again, what the members of the sub-family have in common is that they share the same slope. In general, what members of the linear family have in common is that they have a constant slope. Next we will examine a family of functions that do not have constant slope.

The square function

A square function is one whose equation contains the second power of x, x^2. For example, $f(x) = x^2 + 2x$ - 4 is a member of the square function family, while $g(x) = x^3$ - x^2 is not.

The graph of every square function is a **parabola.** A parabola has a **vertex,** and an **axis of symmetry.** The graph below shows these aspects of the graph of $y = x^2$ - 3.



In chapter 2 you will spend much more time analyzing the graphs of square functions, often referred to as **quadratic** functions. Here, what is important to note is that these functions form a family because they behave in similar ways. For example, because of the shape of a parabola, both ends of the function will have the same behavior.

Example 2. Describe the end behavior of each function:

Table 2.16:

a. $y = x^2$ - 1	b. $y = -x^2 + 1$

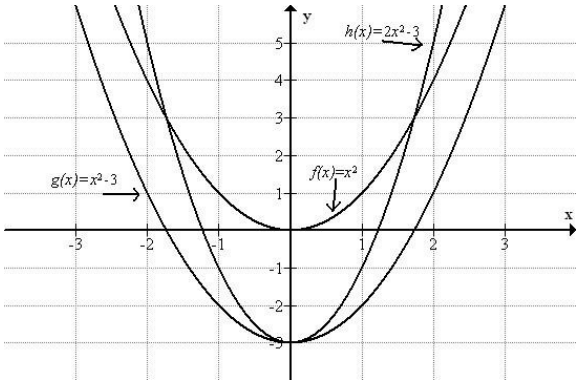
Solution:

a. $y = x^2$ - 1
The graph of this function is a parabola that opens up. Therefore $\lim_{x\to\pm\infty}(x^2 - 1) = \infty$.
b. $y = -x^2 + 1$
The graph of this function is a parabola that opens down. Therefore $\lim_{-x\to\pm\infty}(x^2 - 1) = -\infty$.

All square functions have either a global maximum or minimum. The location of the maximum or minimum is always the vertex of the parabola. Square functions also share behavior in terms of their average rate of change. Consider for example the functions $f(x) = x^2$, $g(x) = x^2$ - 3 , and $g(x) = 2x^2$ - 3. The table below shows the average rate of change (ARC) of each function on several intervals (Note: The ARC of a function on the interval (a, b) is $\frac{f(b)-f(a)}{b-a}$.)

Table 2.17:

Interval	ARC of $f(x)$	ARC of $g(x)$	ARC of $h(x)$
(-1, 0)			
(0, 1)	1	1	2
(1, 2)	3	3	6
(2, 3)	5	5	10

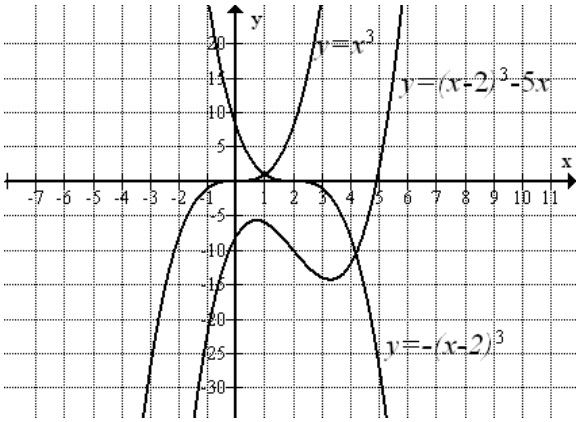


Notice that the average rate of change of $f(x)$ and $g(x)$ is the same on each interval, and the average rate of change of $h(x)$ is twice that of the other two functions. You may also notice that the average rate of change follows a linear pattern: on each interval the rate increases at a constant rate of 2. While linear functions have a constant rate of change, quadratic functions have an average rate of change that follows a linear pattern. (You will examine this pattern in more detail when you study calculus.)

Square functions belong to a larger family referred to as polynomial functions. Next we will examine another member of that larger family.

The cube function

A cube function is one whose highest exponent is x^3 and which does not contain negative or fractional exponents. Much like the square, or quadratic functions, the cube functions can be described by the similar behavior of the graphs. In general, the graphs of cube functions have a particular shape, illustrated by the graph shown here:



As you can see from the graph, each cubic function has a similar shape. However, only some cubic functions will have a relative maximum and minimum. For example, the graph of $y = (x - 2)^3 - 5x$ shown above, has a relative maximum around $x = 0.7$, and a relative minimum around $x = 3.3$. The shape of the cubic graph means that we can predict end behavior: one end will approach ∞, and the other will approach $-\infty$. For example, for both $y = x^3$ and $y = (x - 2)^3 - 5x$ shown above, $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to\infty} f(x) = -\infty$. The third function shown above is $y = -(x - 2)^3$. The end behavior of this function is the opposite of the other two: $\lim_{x\to\infty} f(x) = -\infty$ and $\lim_{x\to-\infty} f(x) = \infty$.

Like the square function, the average rate of change of a cubic function is not constant. The table below shows the average rate of change of y = x^3 on several intervals:

Table 2.18:

Interval	ARC of $y = x^3$
(0, 1)	1
(1, 2)	7
(2, 3)	21
(3, 4)	64

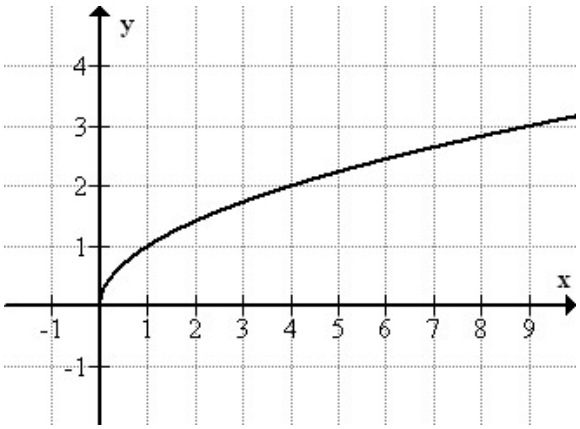
Notice that between each interval, the average rate of change increases faster and faster. When you study calculus, you will learn to model this rate of increase with a quadratic function. What is important to note here is that the cubic function grows faster than an associated quadratic function. For example, $y = x^3$ grows faster than $y = x^2$. The square function and the cube function are part of two larger families: the power functions, and polynomials functions, which you will study in Chapter 2. Now we will turn to several other functions that are slightly more complicated in terms of their behavior and their graphs.

The square root function

The square function and the cube function are, in a sense, well-behaved. They both have a domain of $\mathbb{R}$, and their graphs are smooth curves that follow particular shapes. The square root function is slightly

more complicated. Consider for example, the parent of the family, $y = \sqrt{x}$. The domain of the function is limited to real numbers 0, as the square root of a negative number is not a real number. Similarly, the range of the function is limited to real numbers 0. This may seem confusing if you think of squares having two roots. For example, 9 has two roots: 3, and -3. However, for $y = \sqrt{x}$, we have to define the function value as the **principal root,** which means the positive root.

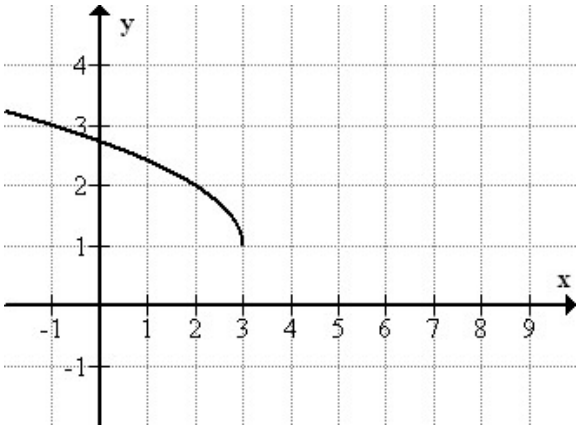
The function $y = \sqrt{x}$ is shown below:



The same kind of limitations of domain and range will exist for any square root function.

Example 3: Graph the function $y = \sqrt{3 - x} + 1$ and state the domain and range of the function.

Solution:



From the graph you can see that the function does not take on any x values above 3. (Why not?) Therefore the domain is limited to real numbers 3. The function's lowest value is 1, so the range is limited to all real numbers 1 .

It is important to note that while the graph of a square root function might look as if it has horizontal asymptote, it does not. The function values will grow without bound (though relatively slowly!).

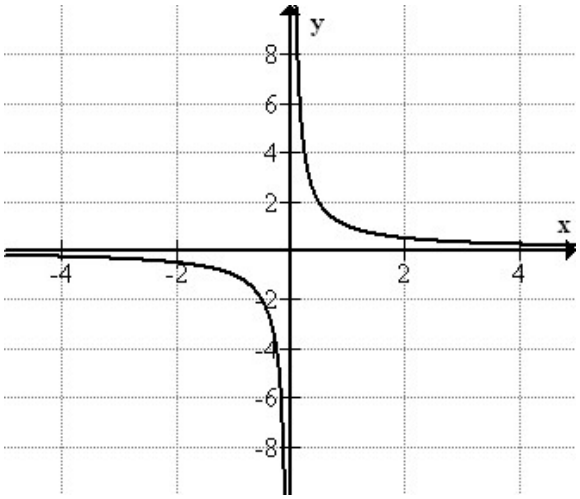
The reciprocal function

An important member of the linear family is the **identity function,** y = x. The graph of this function is a line that passes through the origin, and has a slope of 1. But what does its reciprocal look like?

The function $y = 1/x$ has a graph that is very different from a line. First, the domain cannot include 0, as the fraction (1/0) is undefined. The range also does not include 0, as a fraction can only be zero if the numerator is zero, and the numerator of $y = 1/x$ is always 1.

In order to understand what these limitations mean for the graph, we will consider function values near $x = 0$ and $y = 0$. First, consider very small values of x. For example, consider $x = 0.01$. At the x value, $y = 1/x = (1/0.01) = 100$. Let's choose an x value even closer to 0, such as $x = 0.001$. Now we have $y = 1/x = (1/0.001) = 1000$. As we get closer and closer to $x = 0$, the function values grow without bound. On the other side of the x-axis, if we choose values closer and closer to $x = 0$, the function values will approach $-\infty$. We can see this behavior in the graph as a vertical asymptote: the graph is asymptotic to the y-axis.

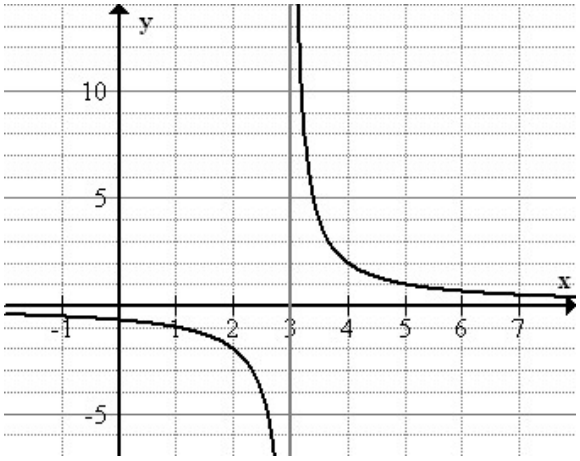
We can also see in the graph that as x approaches $+\infty$ or $-\infty$, the function values approach 0. The exclusion of $y = 0$ from the range means that the function is asymptotic to the x-axis.



The reciprocal function belongs to a larger family of functions called rational functions, which you will examine further in chapter 2. Here we will focus on two key aspects of reciprocal functions: horizontal and vertical asymptotes.

Example 4: Graph the function $f(x) = 2/(x - 3)$ and identify horizontal and vertical asymptotes.

Solution:

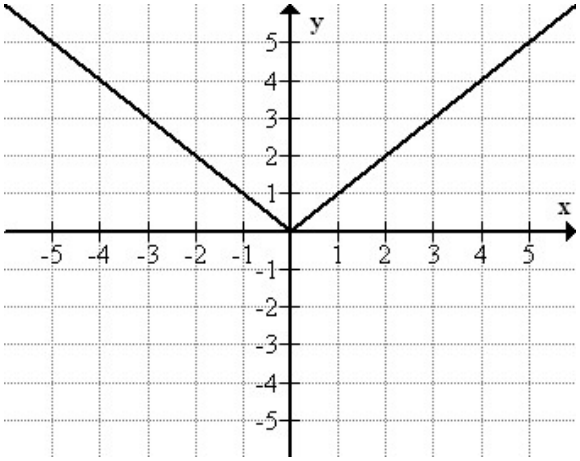


The graph is asymptotic to the x-axis ($y = 0$) and to the line $x = 3$.

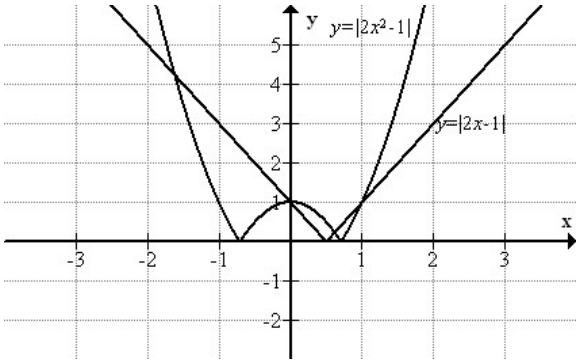
The last three families of functions we have considered so far have all been "smooth" curves. Next we will consider a different kind of function.

Absolute Value Function

Let's first consider the parent of the family, $y = |x|$. Because the **absolute value** of a number is that number's distance from zero, all of the function values will be non-negative. If $x = 0$, then $y = |0| = 0$. If x is positive, then the function value is equal to x. For example, the graph contains the points (1, 1), (2, 2), (3, 3), etc. However, when x is negative, the function value will be the opposite of the number. For example, the graph contains the points (-1, 1), (-2, 2), (-3, 3), etc. As you can see in the graph below, the absolute value function forms a "V" shape.



There are two important things to note about the graph of this kind of function. First, like the square function, the absolute value graph has a vertex and a line of symmetry. For example, the graph of $y = |x|$ has its vertex at (0, 0) and it is symmetric across the y-axis. Second, note that the graph is not curved, but composed of two straight portions. Every absolute value graph will take this shape, as long as the expression inside the absolute value is linear. The graph below shows the graphs of $y = |2x - 1|$ and $y = |2x^2 - 1|$



The graph of $y = |2x - 1|$ makes a "V" shape, much like $y = |x|$. The graph of $y = |2x^2 - 1|$ is curved, and it does not have a vertex, but two "cusps."

Because the graphs of functions such as $y = |x|$ and $y = |2x - 1|$ have this two-sided shape, we can actually define an absolute value function in terms of two lines. This would be an example of a piece-wise defined function.

Piece-wise defined functions

Consider again the function $y = |x|$. For positive x values, the graph resembles the identity function $y = x$. For negative x values, the graph resembles the function $y = -x$. We can express this relationship by defining the absolute value function in two pieces:

$$f(x) = \begin{cases} -x, x < 0 \\ x, x \geq 0 \end{cases}$$

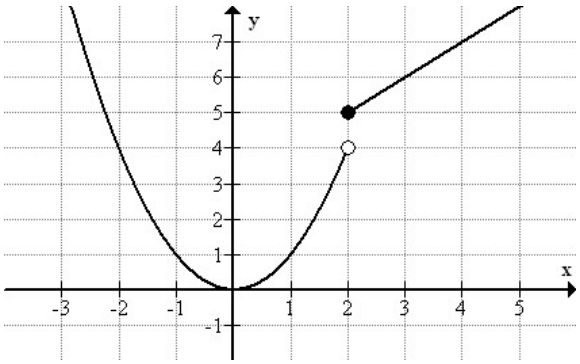
We can read this notation as: the function values are equal to $-x$ if x is negative. The function values are equal to x if x is 0 or positive.

A piece-wise defined function does not have to represent a function that can already be written as a single equation, such as the absolute value function. For example, one "piece" may be from one function family, while another piece is from a different function family.

Example 5: Sketch a graph of the function

$$f(x) = \begin{cases} x^2, x < 2 \\ x + 3, x \geq 2 \end{cases}$$

Solution:



It is important to note that the pieces of a piece-wise defined function may or may not meet up. For example, in the graph of $f(x)$ above, the function value is 4 at $x = -2$, but the piece of the graph that is defined by $x + 3$ is headed to the y value of 1. Therefore the two pieces do not meet.

Lesson Summary

In this lesson we have examined several families of functions: linear, square, cube, square root, reciprocal, absolute value, and piece-wise defined functions. Each family can be characterized by the behavior of the graphs. For example, square functions and absolute value functions have vertices. Linear functions have a constant rate of change, or slope, while others do not. Reciprocal functions have vertical and horizontal asymptotes, while others do not. Finally, piece-wise defined functions are functions that are made up of pieces of any one of these other types of functions. These key aspects of graphs will be useful in the next lesson, as we further explore the graphs of families of functions.

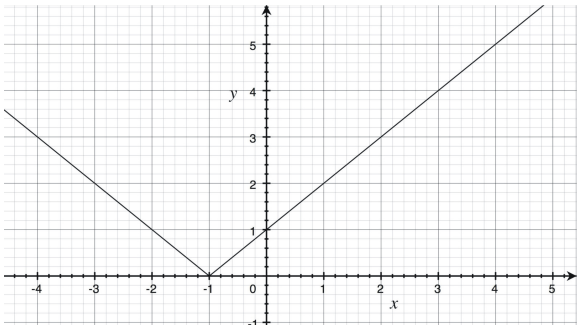
Points to Consider

1. How can you use a graph to determine what family a function belongs to?
2. In general, what families of functions have restrictions on their domains and ranges, and how do these restrictions "show up" in the graphs?
3. What makes a function family a "family"?

Review Questions

1. What is the slope and the y-intercept of the line defined by the equation $2y + 5x = 4$?
2. Consider the function defined by the equation $f(x) = -x^2 + 2x$.a. To what family does this function belong?
   b. State the domain and range of the function.
   c. Describe the end behavior of the function.
3. Sketch a graph of $g(x) =$
4. Describe the end behavior of the function $y = \sqrt{x+4}$.
5. What are the asymptotes of the reciprocal function $y = (3/x) + 4$?
6. Explain how you can identify a graph as a member of the cube family.
7. Give an example of an absolute value function that takes on negative function values.
8. Write the function $f(x) =$
9. What is the average rate of change of the quadratic function $y = (1/2)\ x^2$ on the interval (2, 6)?
10. In the piece-wise defined function shown below, what must the value of a be in order for the two pieces of the functions to meet at $x = 2$?$h(x) = \begin{cases} x + 10, x < 2 \\ ax^2, x \geq 2 \end{cases}$

Review Answers

1. The slope is -5/2, and the y-intercept is (0, 2).
2. a. This is a square, or quadratic function. b. The domain is the set of all real numbers. The range is the set of all real numbers less than or equal to 1.
   c. As x approaches $\pm\infty$, the function values approach $-\infty$.
3. 

The domain of the function is all real numbers. The range of the function is all real numbers greater than or equal to 0.
4. As x approaches infinity, the function values also approach infinity. The domain is limited to values - 4, so the values do not approach negative infinity.
5. The graph of the function is asymptotic to the lines $y = 4$ and $x = 0$.
6. The graph will be a smooth curve, and the ends will have opposite behavior: one end goes up, the other goes down.
7. Example: $f(x) =$
8. $f(x) = \begin{cases} -(x-2), x < 2 \\ x - 2, x \geq 2 \end{cases}$
9. The average rate of change on this interval is (16/4) = 4.
10. a = 3

2.6 Transformations of functions

Learning objectives

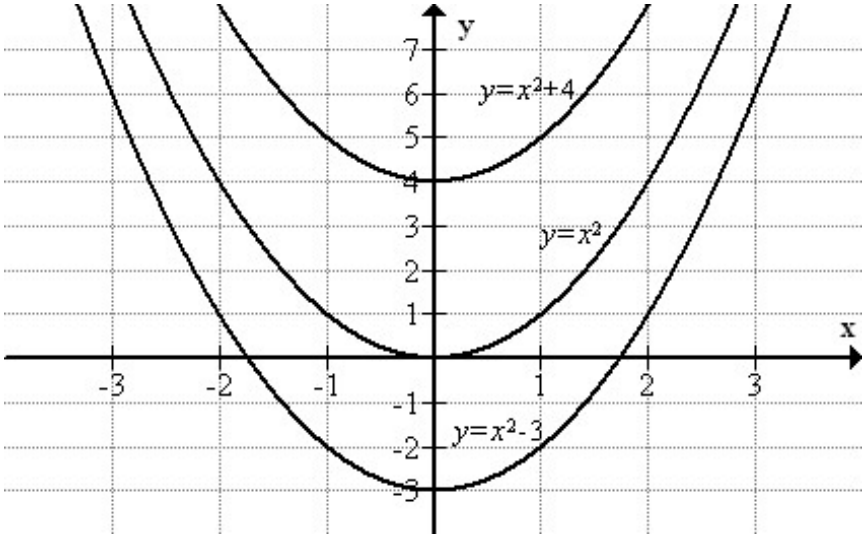
- Graph functions by shifting basic functions.
- Graph functions by stretching or compressing basic functions.
- Graph functions that are reflections of basic functions.
- Combine transformations in order to graph more complicated functions.

Introduction

In the previous lesson we examined the key characteristics of different families of functions. In this lesson we will focus on the relationships among functions within a given family. In particular, we will analyze how a function is related to the parent function in the family (e.g., $y = x^2$, $y = x^3$) so that we can efficiently identify the key characteristics of a particular function and sketch its graph. We will consider seveal kinds of relationships between functions and their parents: shifts, stretches or compressions, and reflections. We will begin with shifts.

Vertical shifts

Consider the graphs of three functions shown below: $y = x^2$, $y = x^2$ - 3, and $y = x^2 + 4$



At first glance, it may seem that the graphs have different widths. For example, it might look like $y = x^2 + 4$, the uppermost of the three parabolas, is thinner than the other two parabolas. However, this is not the case. The parabolas are **congruent**. That is, if we shifted the graph of $y = x^2$ up four units, we would have the exact same graph as $y = x^2 + 4$. Similarly, if we shifted $y = x^2$ down three units, we would have the graph of $y = x^2$ - 3.

Numerically, we can determine that for any value of x, the value of $y = x^2 + 4$ is 4 more than the value of $y = x^2$, and the value of $y = x^2$ - 3 is 3 less than the value of $y = x^2$. The table below shows several function values:

Table 2.19:

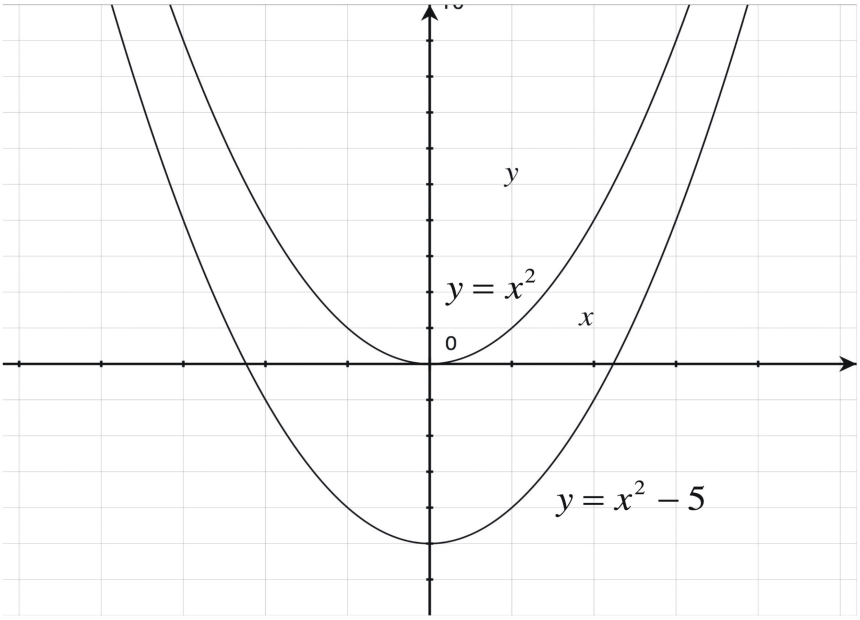
x	$y = x^2$	$y = x^2 + 4$	$y = x^2$ - 3
1	5		
0	0	4	
1	1	5	
2	4	8	1
5	25	29	22
10	100	104	97

We can formally express this relationship using function notation: for any function $f(x)$, the function $g(x) = f(x) + c$ has a graph that is the same as $f(x)$, shifted c units vertically. If c is positive, the graph is shifted up. If c is negative, the graph is shifted down. We can then use this relationship to graph functions by hand.

Example 1. Use the graph of $y = x^2$ to graph the function $y = x^2$ - 5.

Solution:

The graph of $y = x^2$ is a parabola with vertex at (0, 0). The graph of $y = x^2$- 5 is therefore a parabola with vertex (0, -5). To quickly sketch $y = x^2$ - 5, you can sketch several points on $y = x^2$, and then shift them down 5 units.



In sum, when we add or subtract a constant from the equation of a function, the graph of the new function is a vertical shift from the original function. We can use similar reasoning to graph functions that are horizontal shifts of parent graphs.

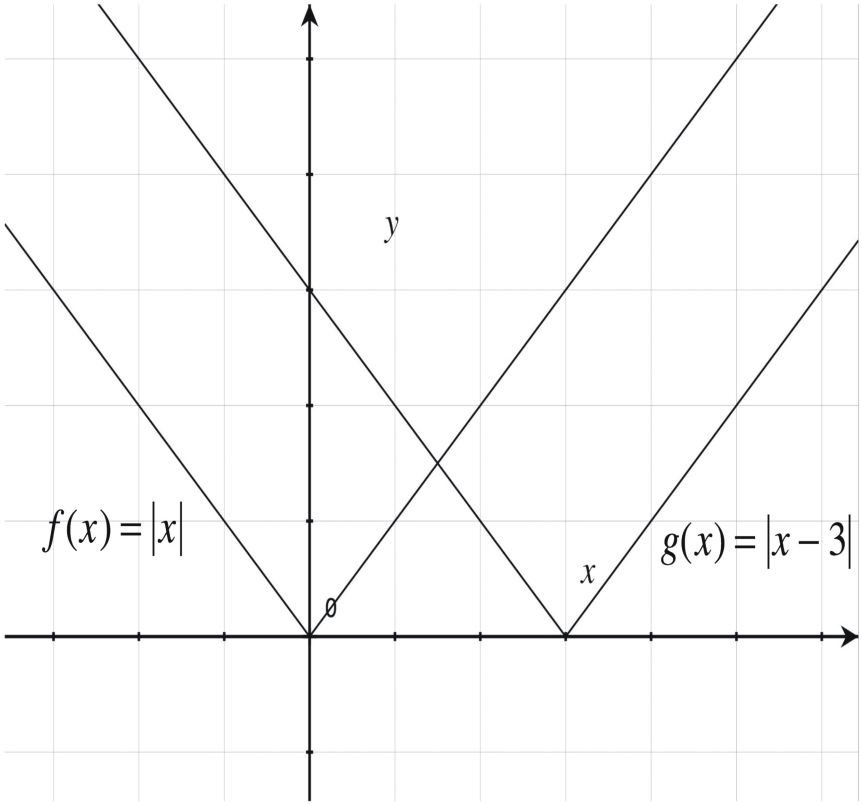
Horizontal shifts

Consider the functions $f(x) = |x|$ and $g(x) = |x - 3|$. From the examples of vertical shifts above, you might think that the graph of $g(x)$ is the graph of f(x), shifted 3 units to the left. However, this is not the case. The graph of $g(x)$ is the graph of $f(x)$, shifted 3 units to the right. The direction of the shift makes sense if we look at specific function values.

Table 2.20:

x	$g(x) = \text{abs}(x - 3)$
0	3
1	2
2	1
3	0
4	1
5	2
6	3

From the table we can see that the vertex of the graph is the point (3, 0). The function values on either side of $x = 3$ are symmetric, and greater than 0.



We can formalize horizontal shifts in the same way we formalized vertical shifts: given a function $f(x)$, and a constant a > 0, the function $g(x) = f(x - a)$ represents a horizontal shift a units to the right from $f(x)$. The function $h(x) = f(x + a)$ represents a horizontal shift a units to the left.

Notice that in the case of a horizontal shift, the constant is part of the "input" of the function. For example, $y = (x + 7)^2$ represents a horizontal shift of the original function $y=x^2$, while $y = x^2 + 7$ represents a vertical shift. It is important to keep this distinction in mind as you interpret equations.

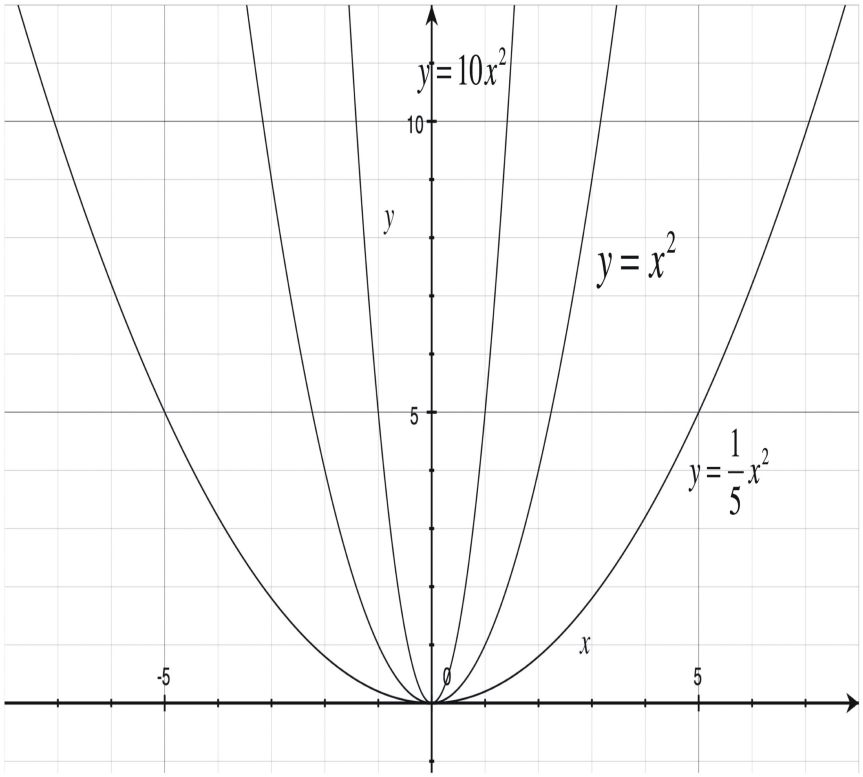
Example 2: What is the relationship between $f(x) = x^2$ and $g(x) = (x - 2)^2 - 6$?

Solution: The graph of $g(x)$ is the graph of $f(x)$, shifted 2 units to the right, and down 6 units.

In sum, adding or subtracting constants within the equation of a function creates a vertical or horizontal shift in the graph. When the constant is "inside" the parentheses, the shift will be horizontal. Now we will consider the results of multiplying a function or the input of a function (x) by a coefficient.

Stretching and compressing graphs

If we multiply a function by a coefficient, the graph of the function will be stretched or compressed. Consider the graph of three parabolas below.



The widest parabola is the function $y = (1/5)\ x^2$. The middle parabola is the function $y = x^2$. The thinnest parabola is the function $y = 10x^2$. From these graphs, we can see that multiplying by different coefficients has different effects on the graph of $y = x^2$. The graph of $y = 10x^2$ represents a **vertical stretch** of the graph of $y = x^2$. Every function value of $y = 10x^2$ is 10 times the equivalent function value of $y = x^2$. For example, $y = x^2$ contains the point (2, 4), while y = $10x^2$ contains the point (2, 40).

The coefficient of 1/5 has an opposite effect. We describe this as a vertical compression. Every function value of $y = (1/5)\ x^2$ is 1/5 the value of $y = x^2$. For example, $y = x^2$ contains the point (10, 100), while $y = (1/2)\ x^2$ contains the point (10, 20).

We can formalize vertical stretching and compressing as follows:

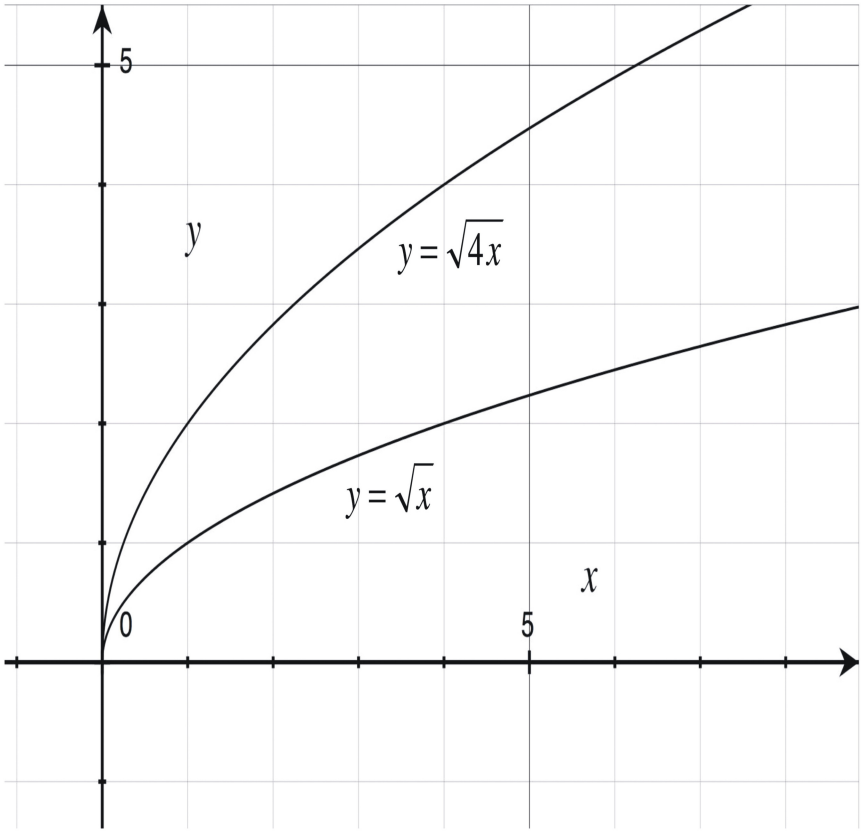
- A function $g(x) = cf(x)$ represents a vertical stretch if $c > 1$.
- A function $h(x) = cf(x)$ represents a vertical compression if $0 < c < 1$.

Notice that the graph of $y = (1/5)\ x^2$ is wider than the other two parabolas. While 1/5 has the effect of compressing the graph of $y = x^2$ vertically, we can also describe the graph as having a **horizontal stretch.** Formally, we characterize a transformation as a horizontal stretch or compression if the function is of the form $g(x) = f(cx)$. That is, instead of multiplying the function by a coefficient, the variable x is multiplied by c. However, if we analyze a quadratic equation, we can see why a vertical compression looks like a horizontal stretch. (The parabola gets shorter/wider, or taller/thinner.)

Consider for example the function $y = (3x)^2$. We have multiplied x by 3. This should affect the graph horizontally. However, if we simplify the equation, we get $y = 9x^2$. Therefore the graph if this parabola will be taller/thinner than $y = x^2$. Multiplying x by a number greater than 1 creates a horizontal compression, which looks like a vertical stretch.

Now consider the function $y = ((1/2)x)^2$. If we simplify this equation, we get $y = (1/4)\ x^2$. Therefore multiplying x by a number between 0 and 1 creates a horizontal stretch, which looks like a vertical compression. That is, the parabola will be shorter/wider.

We can see similar behavior in other functions. For example, consider the functions $y = \sqrt{x}$, $y = \sqrt{4x}$, and $y = 2\sqrt{x}$, shown in the graph below. If we simplify the second equation, we get $y = \sqrt{4x} = \sqrt{4}\sqrt{x} = 2\sqrt{x}$. Therefore the second and third equations are actually the same function, and we can describe the function as a vertical stretch or as a horizontal compression.



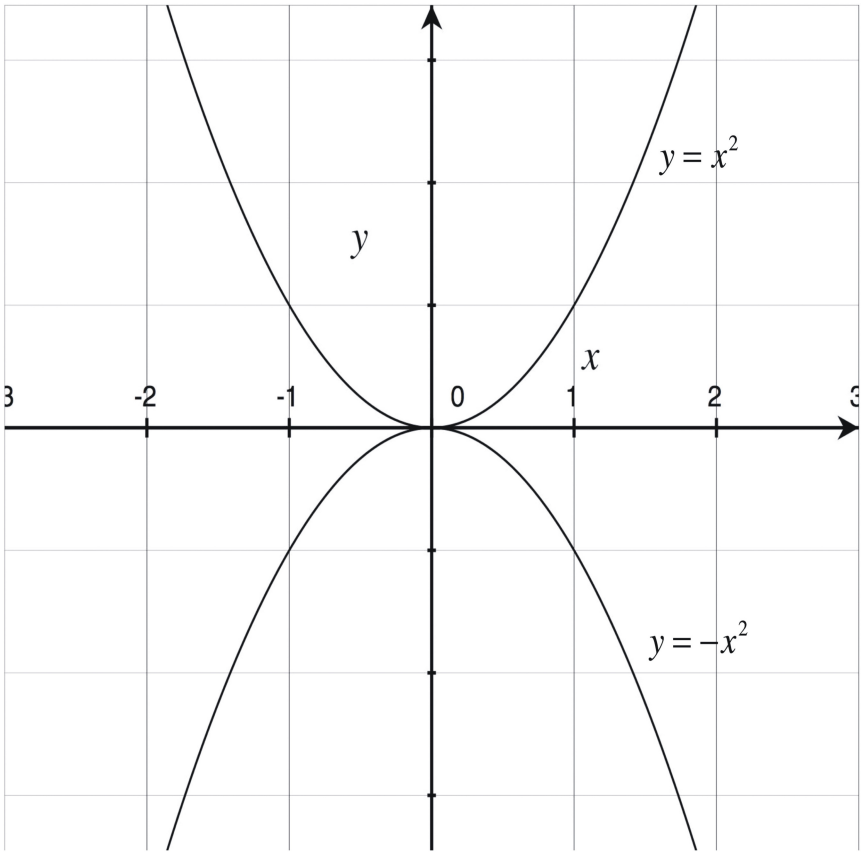
Given a function $f(x)$, we can formalize compressing and stretching the graph of $f(x)$ as follows:

- A function $g(x)$ represents a vertical stretch of $f(x)$ if $g(x) = cf(x)$ and $c > 1$.
- A function $g(x)$ represents a vertical compression of $f(x)$ if $g(x) = cf(x)$ and $0 < c < 1$.
- A function h(x) represents a horizontal compression of $f(x)$ if $h(x) = f(cx)$ and $c > 1$.
- A function $h(x)$ represents a horizontal stretch of $f(x)$ if $h(x) = f(cx)$ $0 < c < 1$.

Notice that a vertical compression or a horizontal stretch occurs when the coefficient is a number between 0 and 1. Now we will consider the effects of a coefficient less than 0.

Reflecting graphs over the y-axis and x-axis.

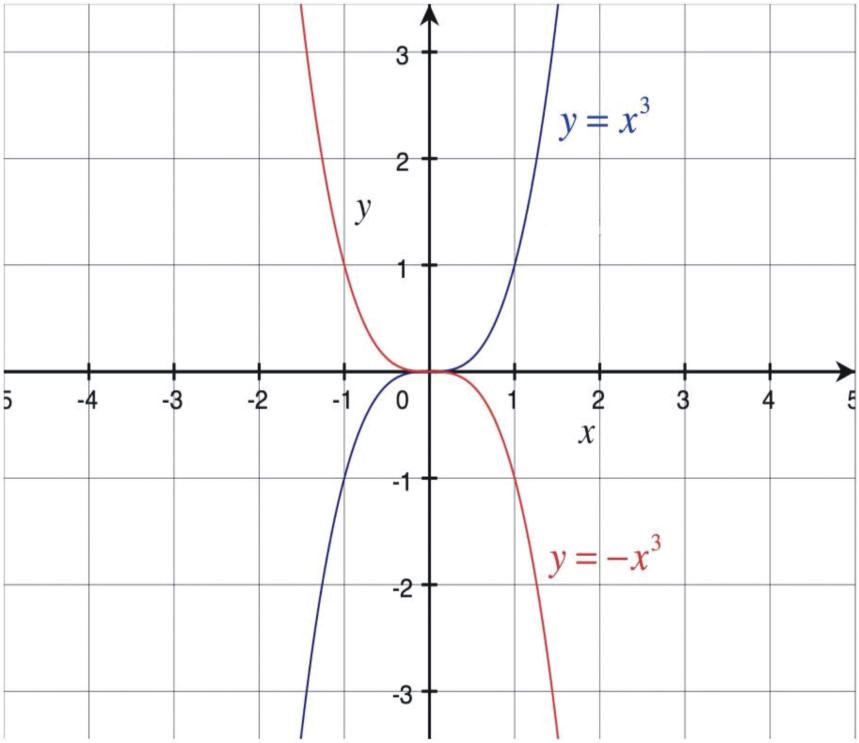
Consider the graphs of the functions $y = x^2$ and $y = -x^2$, shown below.



The graph of $y = -x^2$ represents a **reflection** of $y = x^2$, over the x-axis. That is, every function value of $y = -x^2$ is the negative of a function value of $y = x^2$. In general, $g(x) = -f(x)$ has a graph that is the graph of $f(x)$, reflected over the x-axis.

Example 3: Sketch a graph of $y = x^3$ and $y = -x^3$ on the same axes.

Solution:



At first the two functions might look like two parabolas. If you graph by hand, or if you set your calculator to sequential mode (and not simultaneous), you can see that the graph of $y = -x^3$ is in fact a reflection of $y = x^3$ over the x-axis.

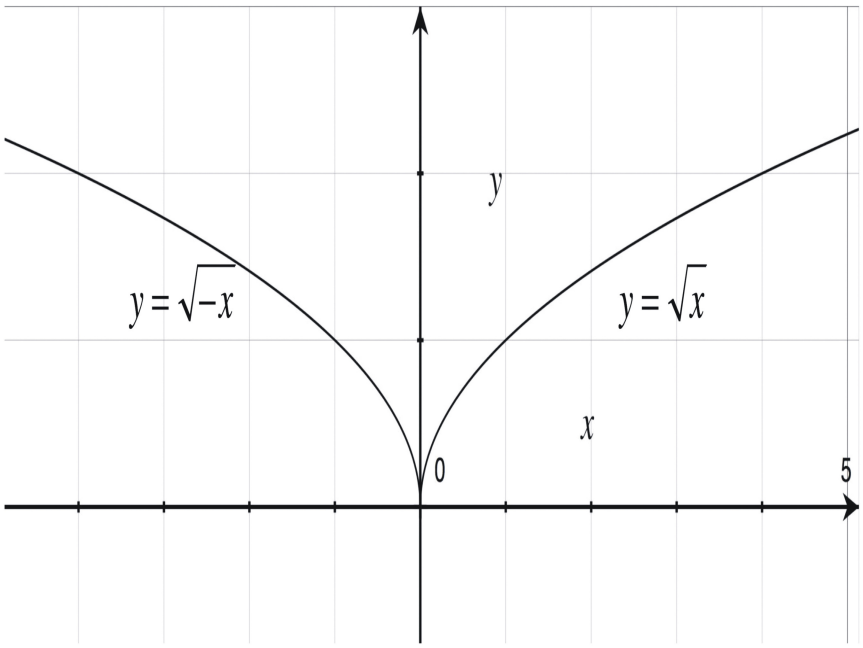
However, if you look at the graph, you can see that it is a reflection over the y-axis as well. This is the case because in order to obtain a reflection over the y-axis, we negate x. That is, $h(x) = f(-x)$ is a reflection of $f(x)$ over the y-axis. For the function $y = x^3$, $h(x) = (-x)^3 = (-x)\ (-x)\ (-x) = -x^3$. This is the same function as the one we have already graphed.

It is important to note that this is a special case. The graph of $y = x^2$ is also a special case. If we want to reflect $y = x^2$ over the y-axis, we will just get the same graph! This can be explained algebraically: $y = (-x)^2 = (-x)\ (-x) = x^2$.

Now let's consider a different function.

Example 4: Graph the functions $y = \sqrt{x}$ and $y = \sqrt{-x}$.

Solution:



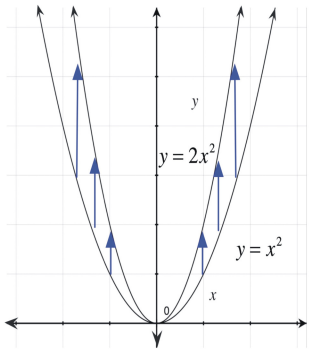
The equation $y = \sqrt{-x}$ might look confusing because of the -x under the square root. It is important to keep in mind that - x means the *opposite* of x. Therefore the domain of this function is restricted to values 0. For example, if x = - 4, $y = \sqrt{-(-4)} = \sqrt{4} = 2$. It is this domain, which includes all real numbers not in the domain of $y = \sqrt{x}$ plus zero, that gives us a graph that is a reflection over the y-axis.

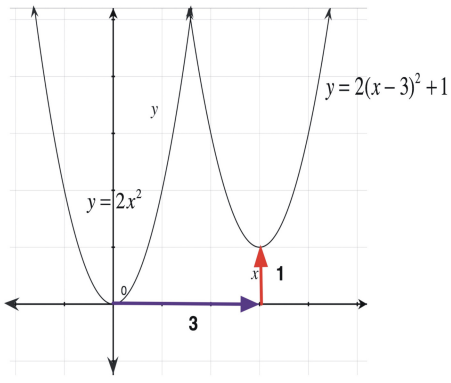
In sum, a graph represents a reflection over the x-axis if the function has been negated (i.e. the y has been negated if we think of $y = f(x)$). The graph represents a reflection over the y-axis if the variable x has been negated. Now that we have considered shifts, stretches/compressions, and reflections, we will look at graphs that combine these transformations.

Combining transformations

Consider the equation $y = 2(x - 3)^2 + 1$. We can compare the graph of this function to the graph of the parent $y = x^2$: the graph represents a vertical stretch by a factor of 2, a horizontal shift 3 units to the right, and a vertical shift of 1 unit.

We can use this relationship to graph the function $y = 2(x - 3)^2 + 1$. You can start by sketching $y = x^2$ or $y = 2x^2$. Then you can shift the graph 3 units to the right, and up 1 unit.





It is slightly more complicated to graph functions that represent reflections of the parent graphs. Next we will examine reflections combined with other transformations.

Example 5. Graph each function using your knowledge of the function $y = |x|$ and your knowledge of transformations.

a. $f(x) = -|x| + 3$
b. $g(x) = |-x + 3|$

Solution:

a. $f(x) = -|x| + 3$

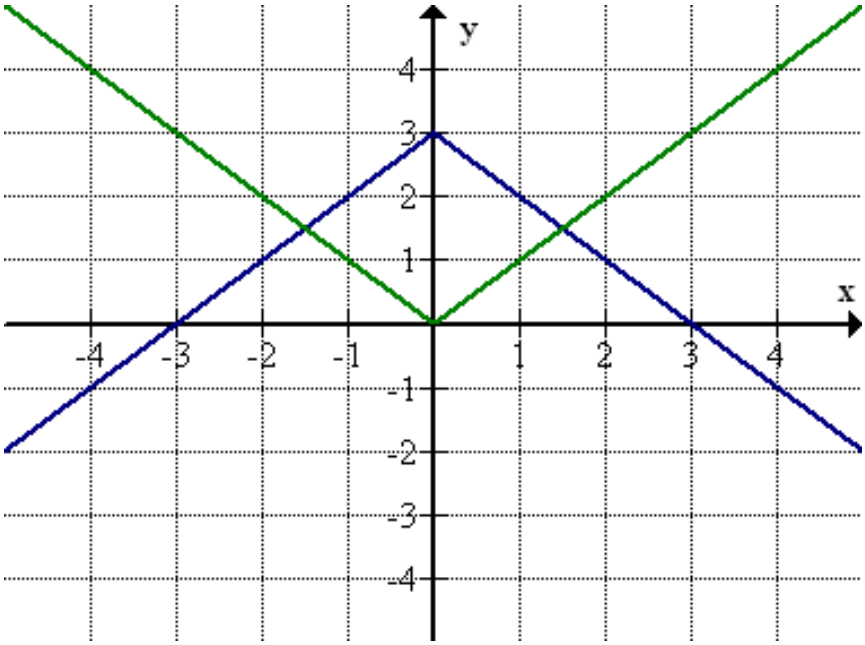
The parent graph of this function is the graph of $y = |x|$, reflected over the x-axis, and shifted up 3 units. The question is: which transformation do you perform first?

We can answer this question if we consider a few key function values. The table below shows several function values for $f(x) = -|x| + 3$:

Table 2.21:

x	$f(x) = -\text{ abs}(x) + 3$
-3	-abs(-3) + 3 = -(+3) + 3 = -3 + 3 = 0
-2	-abs(-2) + 3 = -(+2) + 3 = -2 + 3 = 1
-1	2
0	3
1	2
3	0

From the function values in the table we can see that the function increases until a vertex at (0, 3), and then it decreases again. This tells us that we can obtain the graph if we first reflect $y = |x|$ over the x-axis (turn the "v" upside down), and then shift the graph up 3 units.



We can also justify this ordering of the transformations of we think about the order of operations. To find any function value we take an x value, find its absolute value, find the negative of that number, and then add 3. This is the same as the order of the transformation: reflection comes before shifting up.

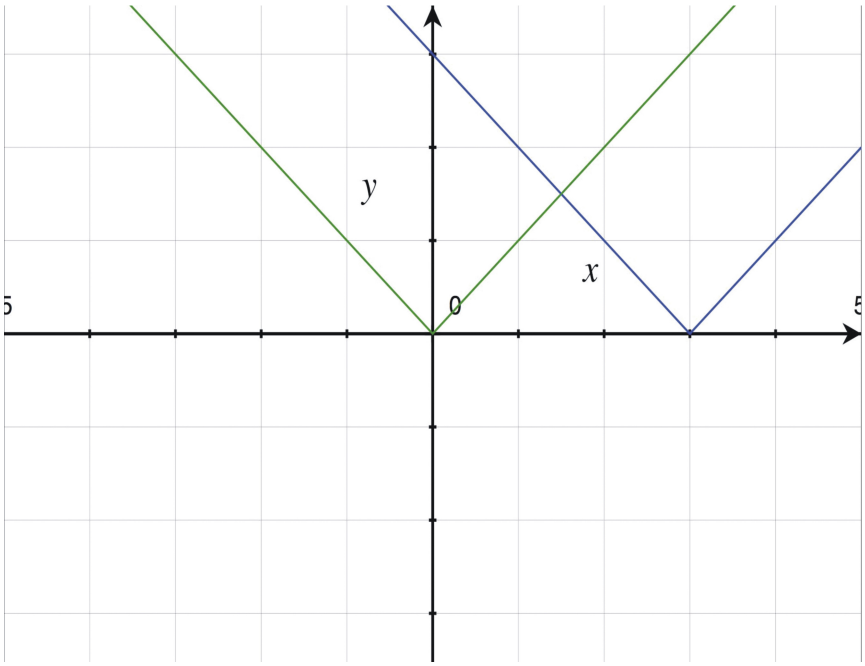
b. $g(x) = |\text{-}x + 3|$

This function represents a horizontal shift of $y = |x|$, and a reflection over the x-axis. Before graphing, consider a few function values:

Table 2.22:

x	$g(x) = abs(-x + 3)$
abs(-(-3) + 3) = abs(3 + 3) = abs(6) = 6	
abs(-(-2) + 3) = (2 + 3) = abs(5) = 5	
4	
0	3
1	2
2	1
3	0
4	1
5	2
6	3

From the values in the table, we can see that the vertex of the graph is at (3, 0). The graph is shown below.



The graph looks the same as the graph of $y = |x - 3|$. This is the case because $y = |\text{-}x + 3| = |\text{-}(x - 3)|$, and because $|\text{-}\ a| = |a|$ for all values of a, then $|\text{-}(x - 3)| = |x - 3|$. So the original function is equal to $|x - 3|$. We can still think of this graph as a reflection: if we reflect $y = |x|$ over the x-axis, the graph remains the same, as it is symmetric over the x-axis. Then we shift the graph 3 units to the right. What is important to note here is that in order to "read" the equation as a horizontal shift, the entire expression inside the function (in this case, inside the absolute value) must be negated.

Lesson Summary

In this lesson we have explored transformations of functions. We have related different functions to their parent graphs, in terms of vertical and horizontal shifts, vertical and horizontal stretches and compressions, and reflections over the y-axis and the x-axis. By analyzing the functions, we can relate the graphs of functions to the graphs of their parent functions, and we can use these relationships to sketch graphs. The issue of "inside" and "outside", raised throughout the lesson, and seen specifically in example 5b, will be examined more closely in the next lesson.

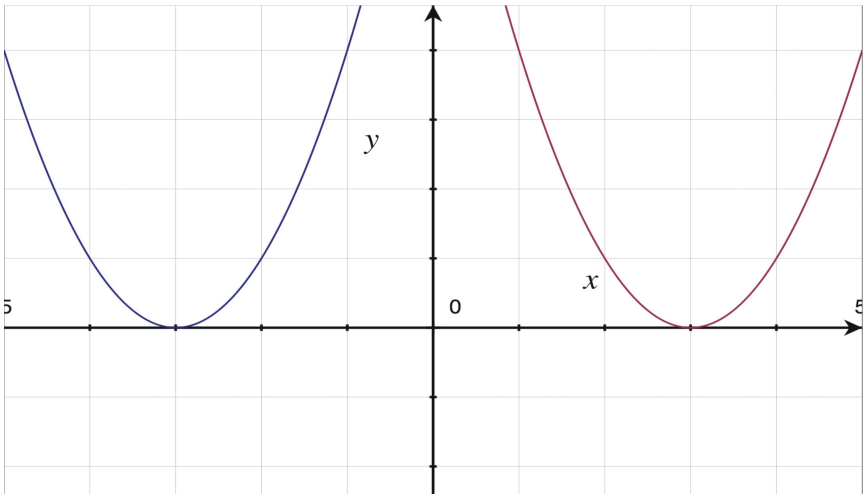
Points to Consider

1. How can you analyze an equation and determine if the graph represents a vertical or horizontal shift?
2. How can you analyze an equation and determine if the graph represents a stretch or compression?
3. How is it that different equations might yield the same graphs?

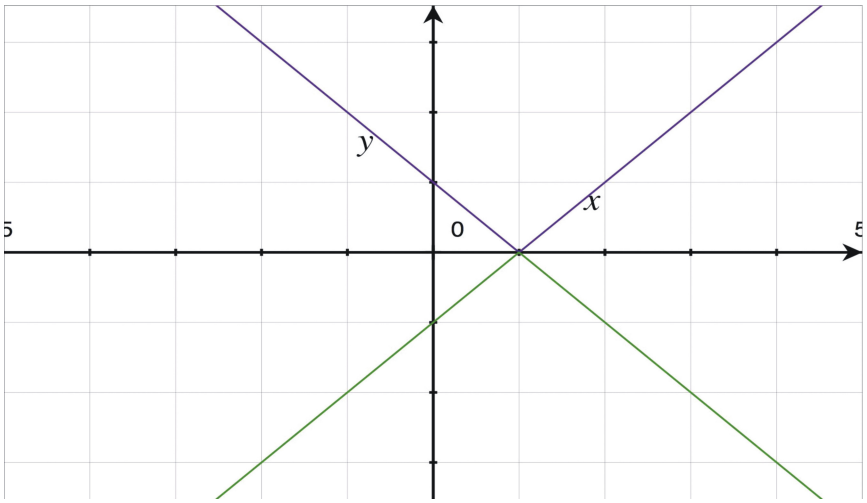
Review Questions

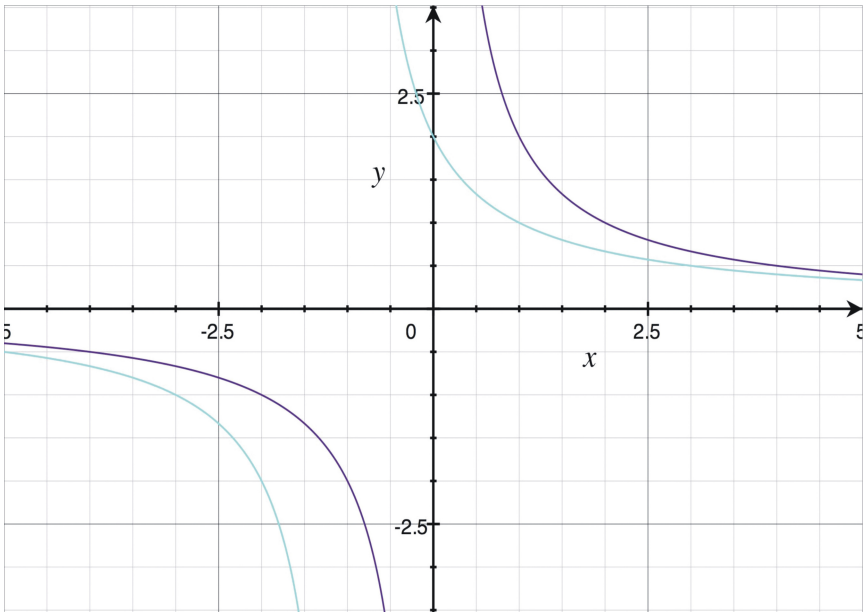
1. How is the graph of $f(x) = x^3 + 5$ related to the graph of $g(x) = x^3$?
2. How is the graph of $f(x) = (x - 3)^2$ related to the graph of $f(x) = (x + 3)^2$? Sketch these functions together.
3. Write an equation that represents a vertical stretch of the graph of $y = \sqrt{x}$ by a factor of 5.
4. Describe the function $f(x) = (5x)^2$ as both a stretch and a compression of $y = x^2$.
5. Graph the function $g(x) =$
6. Describe the relationship between the graphs of $f(x) = 4(x + 8)^3 - 3$ and $g(x) = x^3$.
7. Graph the functions $f(x) = \frac{2}{x}$ and $h(x) = \frac{2}{x+1}$ on the same graph. State the asymptotes of both functions.
8. Graph the functions $f(x) = 3x + 1$ and $f(-x)$ on the same grid. What is the equation of $f(-x)$?
9. Explain how a function can be its own reflection over an axis.
10. Give an example of two equations that can be represented by the same graph.

Review Answers

1. The graph of f is the graph of g, shifted up 5 units.
2. 

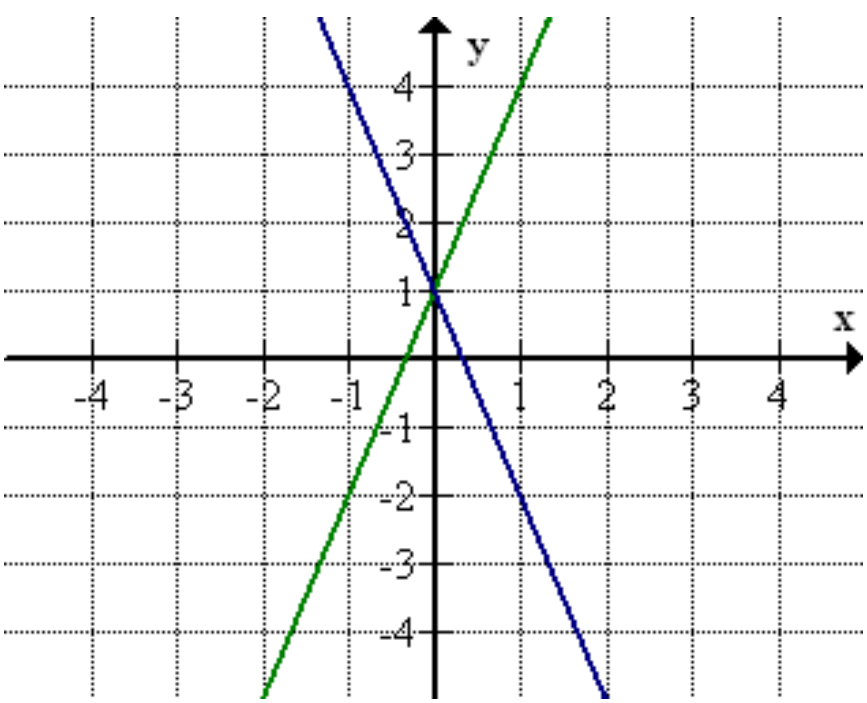
The graph of f is the graph of g, shifted 6 units to the right.

3. $y = 5\sqrt{x}$
4. This function is a horizontal compression by a factor of 5, or a vertical stretch by a factor of 25.
5. 

$-g(x) = -$
6. The graph of f is the graph of g, stretched vertically by a factor of 4, shifted 8 units left, and shifted 3 units down.
7. 

Asymptotes of f: y=0, x=0

Asymptotes of h: y=0, x = - 1

8. 

$f(-x) = 3(-x) + 1 = -3x + 1$

9. Answers will vary. Example: Any graph that is symmetric across the x-axis will be its own reflection over the x-axis. The sides will be "reflected" into each other, or they will "switch" positions.
10. Answers will vary Example: $y = (-x)^2$ and $y = x^2$

2.7 Operations and Compositions

Learning objectives

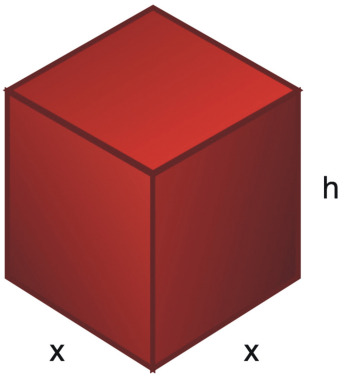
- Perform operations on functions, including adding, subtracting, multiplying, and dividing functions.
- Determine the characteristics of the graph that results from these operations.
- Compose and decompose functions.
- Determine the nature of the graph of composed functions.

Introduction

In the previous two lessons we examined the graphical properties of different families of functions. In this lesson we will work with functions algebraically. Just as we can perform operations on numbers, so too can we perform operations on functions. We will work with basic operations – addition, subtraction, multiplication, and division – as well as a kind of operation called "composition" that combines two functions in a particular way. We will then interpret these operations in terms of the characteristics of the resulting graphs. We will begin with adding and subtracting functions.

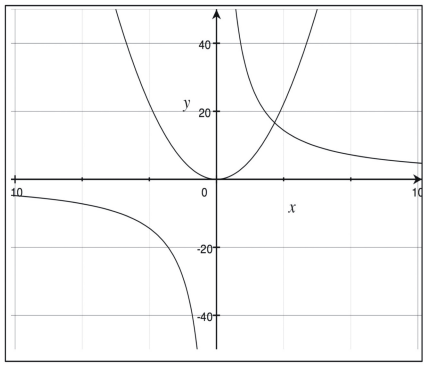
Sums and Differences of Functions

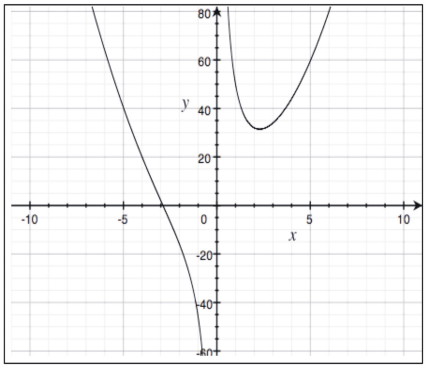
Recall that in lesson 2 we looked at examples of functions that modeled different situations. One of the functions we examined represented the surface area of a box whose base was a square of side-length x.



To express the surface area of the box as a function of x, we began with the equation $S = 4xh + 2x^2$. The h in the equation represented the height of the box. Given that the box has a fixed volume of 12, we were able to rewrite the equation as $s(x) = 48/x + 2x^2$.

Notice that the equation has two terms. The first term, $48/x$, came from the first term of the original equation, $4xh$. This term represents the combined area of the four lateral faces of the box. The other term in the equation, $2x^2$, represents the combined area of the top and bottom of the box. Therefore we can think of the function $S(x)$ as the sum of two other functions: $L(x) = 48/x$ and $E(x) = 2x^2$. Notice that $L(x)$ is a reciprocal function. It is asymptotic to $x = 0$ and $y = 0$. $E(x)$, on the other hand, is a quadratic function. When we add the functions together, we get a new type of graph that resembles both the graphs of $L(x)$ and $E(x)$:





The graph on the right is $S(x)$. The right portion of $S(x)$ resembles the parabola $E(x)$, but is important to keep in mind that this section of $S(x)$ is *not* a parabola. This point is particularly clear if we note that this portion of $S(x)$ is asymptotic to the y-axis. The left portion of $S(x)$ resembles the left side of $L(x)$, as both functions are asymptotic to the negative y-axis.

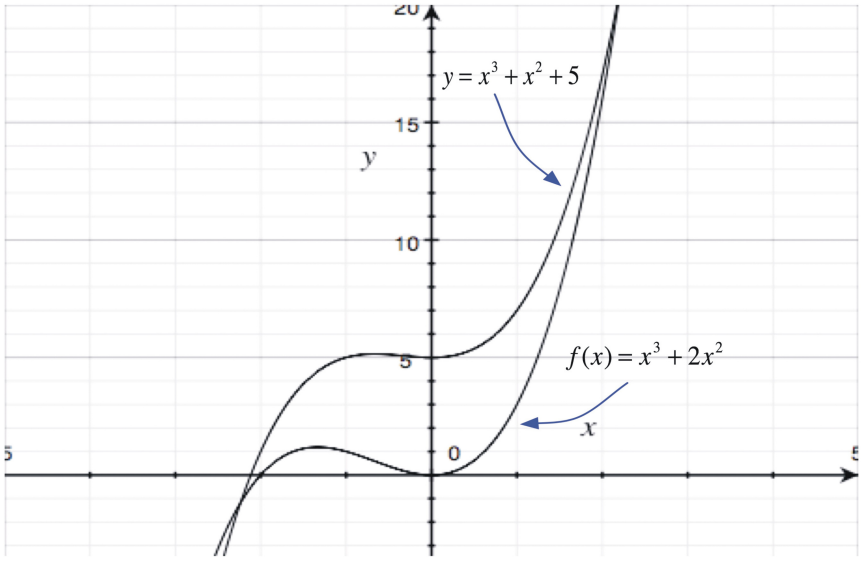
There are two points to be stressed here: first, that we can add functions together, and second, that the resulting sum may be a different kind of function from the original two.

The sum or difference of a function is more likely to resemble the original two functions if they are from the same family. For example, if $f(x) = 3x + 1$ and $g(x) = 4x - 5$, then $f + g = f(x) + g(x) = 3x + 1 + 4x - 5 = 7x - 4$. So the sum function is also a member of the linear family.

Example 1: If $f(x) = x^3 + 2x^2$ and $g(x) = x^2 - 5$, what is $f - g$? What does the graph look like?

Solution: The difference is: $f - g = x^3 + 2x^2 - (x^2 - 5) = x^3 + 2x^2 - x^2 + 5 = x^3 + x^2 + 5$

The graph of the new function, along with $f(x)$, is shown here:



Because $f(x)$ and the new function $y = x^3 + x^2 + 5$ are both members of the cubic family, they have similar shapes.

In sum, when we add or subtract functions, the resulting sum or difference function may be in the same family as one or both of the original functions, or it may be a different type of function. Next we will multiply and divide functions, which are operations that will result in different kinds of functions.

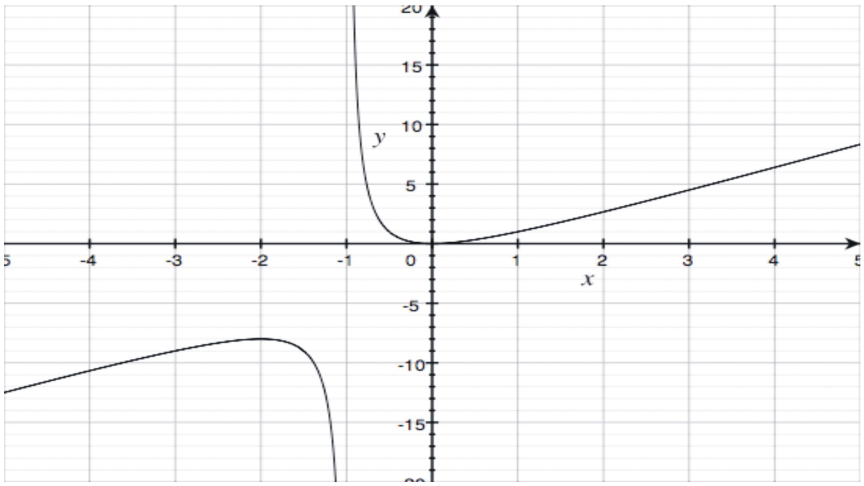
Multiplying and dividing functions

Consider the functions $f(x) = 2x^2$ and $g(x) = x + 1$. If we multiply these functions together, we get a new function: $f \cdot g = f(x) \cdot g(x) = h(x) = 2x^2 \ (x + 1) = 2x^3 + 2x^2$. The product $h(x)$ is a cubic function. If we multiply any linear function by a quadratic, the resulting product function will be a cubic function. However, if we divide, the quotient will be a member of the rational function family.

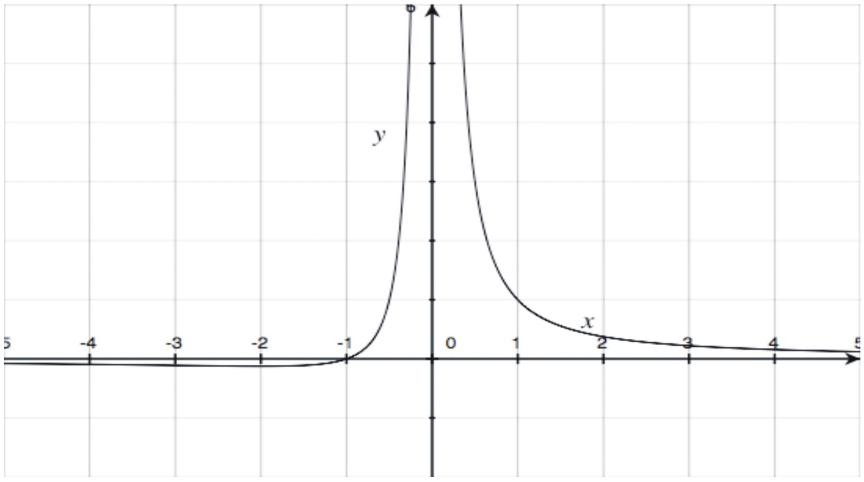
Example 2. Given $f(x) = 2x^2$ and $g(x) = x + 1$, find $r(x) = f(x)/\ g(x)$ and $t(x) = g(x)/f(x)$.

Solution:

$r(x) = f(x)/g(x) = 2x^2/(x + 1)$. This is a rational function. Like some of the functions we studied in lesson 5, this function does not have a horizontal asymptote. It does, however, have a vertical asymptote at $x = -1$, as the domain excludes $x = -1$.



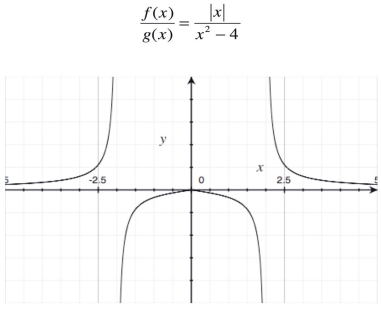
$t(x) = g(x)/\ f(x) = (x + 1)/2x^2$. This is also a rational function. This function has a horizontal asymptote at $y = 0$ (the x-axis), and a vertical asymptote at $x = 0$ (the y-axis).

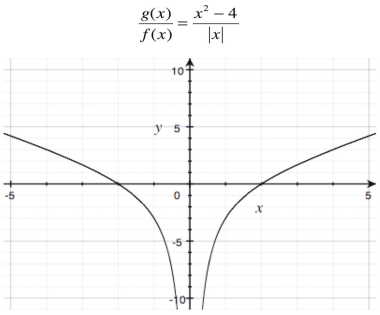


Notice that the graph of this function crosses its asymptote at (-1, 0), but then as x approaches $-\infty$, the function values approach 0.

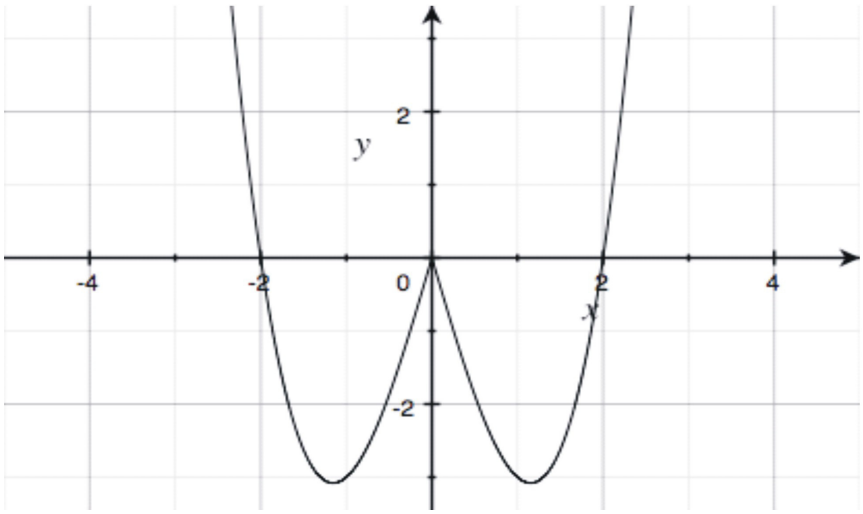
In general, if we multiply linear and polynomial functions (quadratics, cubics, and other such functions with higher exponents, such as $y = x^4 + 3x^2 + 2$), we will obtain other polynomial functions. If we divide these kinds of functions, we will obtain other polynomial functions, or rational functions. If we multiply and divide other types of functions, the resulting functions may be more complicated. For example, consider the functions $f(x) = |x|$ and $g(x) = x^2$ -4.

The graphs of $f(x)/g(x) = |x|/(x^2 - 4)$ and $g(x)/f(x) = (x^2 - 4)/\ |x|$ are shown below:





The graphs of these two functions are not unlike the rational functions you will study in chapter 2. Now consider the function $f(x) \cdot\ g(x)$:



This function does not resemble the rational functions. It does, however, resemble the graph of $y = |2x^2 - 1|$, which we saw in Lesson 5. The point to be stressed here is that multiplying and dividing functions may result in very different functions that would be challenging to graph with out the aid of technology!

So far we have examined what happens to functions when we perform the four basic operations of addition, subtraction, multiplication, and division. Now we will consider a different kind of operation.

Composition of Functions

Functions are often described in terms of "input" and "output." For example, consider the function $f(x) = 2x + 3$. When we input an x value, we output a y value, or a function value. We find the output by taking the input x, multiplying by 2, and adding 3. We can do this for any value of x. Now consider a second function $g(x) = 5x$. For this function too, we can take an x value, input the x into $g(x)$, and obtain an output. For example, if $x = 4$, $g(4) = 5(4) = 20$. What happens if we then take the output of 20 and use it as the input of f?

Then we have $f(20) = 2(20) + 3 = 43$. The table below shows several examples of this same process:

Table 2.23:

x	Output from g	Output from f
2	10	23
3	15	33
4	20	43
5	25	53

Examining the values in the table, we can see a pattern: all of the final output values from f are 3 more than 10 times the initial input. We have created a new function in which $g(x) = 5x$ is the input:

$h(x) = f(5x) = 2(5x) + 3 = 10x + 3$

When we input one function into another, we call this the **composition** of the two functions. Formally, we write the composed function as $f(g(x)) = 10x + 3$ or $(f \circ g)x = 10x + 3$

Example 3: Find $f(g(x))$ and $g(f(x))$:

Table 2.24:

a. $f(x) = 3x + 1$ and $g(x) = x^2$	b. $f(x) = 2x + 4$ and $g(x) = (1/2)x$-2

Solution:

a. $f(x) = 3x + 1$ and $g(x) = x^2$

$f(g(x)) = f(x^2) = 3(x^2) + 1 = 3x^2 + 1$
$g(f(x)) = g(3x + 1) = (3x + 1)^2 = 9x^2 + 6x + 1$
In both cases, the resulting function is quadratic.

b. $f(x) = 2x + 4$ and $g(x) = (1/2)\ x$ - 2

$f(g(x)) = 2((1/2)x$ -2) + 4 = (2/2)x - 4 + 4 = (2/2)$x = x$

$g(f(x)) = g(2x$+4) = (1/2)(2x+4)-2 = x+2=2=x.
In this case, the composites were equal to each other, and they both equal x, the original input into the function. This means that there is a special relationship between these two functions. We will examine this relationship in Chapter 3. It is important to note, however, that $f(g(x)$ is not necessarily equal to $g(f(x))$.

When we compose functions, we are combining two (or more) functions by inputting the output of one function into another. We can also *decompose* a function. Consider the function $f(x)=(2x+1)^2$. We can decompose this function into an "inside" and an "outside" function. For example, we can construct $f(x) = (2x+1)^2$ with a linear function and a quadratic function. If $g(x)=x^2$ and $h(x)=(2x+1)$, then $f(x) = g(h(x))$. The linear function $h(x) = (2x+1)$ is the inside function, and the quadratic function $g(x) = x^2$ is the outside function.

Example 4: Decompose the function $f(x)=(3x\text{-}1)^2\text{-}5$ into a quadratic function $g(x)$ and a linear function $h(x)$.

Solution: Let $h(x)=3x$-1 and $g(x)=x^2$-5. Then $f(x)=g(h(x))$ because $g(h(x))=g(3x\text{-}1)=(3x\text{-}1)^2\text{-}5$.

The decomposition of a function is not necessarily unique. For example, there are many ways that we could express a linear function as the composition of other linear functions. (You will be asked to find decompositions of linear functions in the review questions.)

Lesson Summary

In this lesson we have performed operations on functions and analyzed the resulting graphs. When we add or subtract functions, the resulting function may or may not be a member of the same family as the original functions. When we multiply linear and polynomial functions, the result will be a polynomial function. When we divide linear and polynomial functions, the result will be a polynomial or a rational function. If we perform operations on non-polynomial functions such as the absolute value function, the resulting function is likely to be somewhat complicated. Using a graphing calculator or graphing software can help you see the features of these graphs.

In this lesson we have also composed and decomposed functions. When we compose functions, we treat one function as the input of another. When we decompose functions, we identify functions that could be composed to produce a more complicated function. We can also use composition of functions to model situations in which there are multiple "inputs." This kind of model, as well as other types of models, will be the focus of the next lesson.

Points to Consider

1. What is the difference between performing operations on functions and composing functions?
2. In what situation will a composed function be a member of the same family as one of the components?

Review Questions

For questions 1 – 4, let $f(x)$=4x-5 and $g(x)$=(1/2)x^2 + x

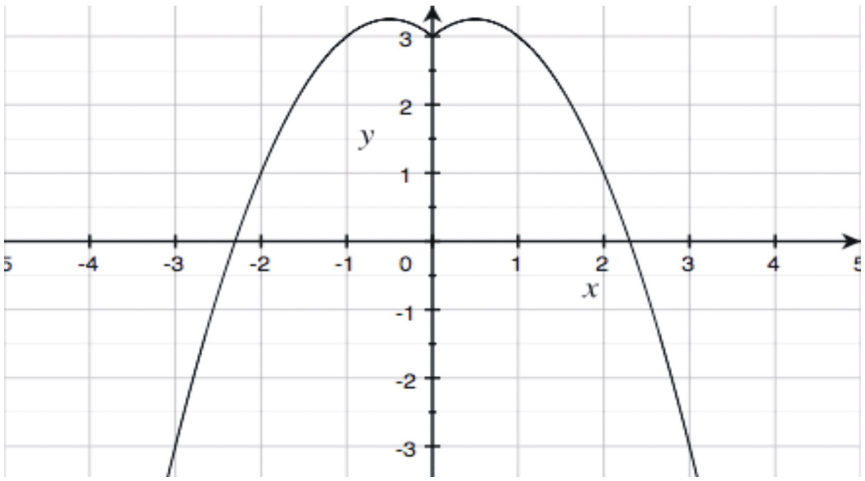
1. $f(x)+g(x)$=?
2. $g(x)$-$f(x)$=?
3. $f(x) \cdot g(x)$=?
4. Write an equation $h(x)= f(x)/g(x)$ and state the domain of $h(x)$.
5. Let $f(x) = x^3$ and $g(x)$=x+8. a. Find $f(g(x))$
   b. Find $g(f(x))$
   c. Are the composed functions equal?
6. How are the graphs of the cubic functions in problem 5 related?
7. Let $f(x)$=2x-5 and $g(x)$=5x-2 . Find $f(g(f(x)))$.
8. Write the function $t(x)$=$(3x+1)^4$ as the composition of two functions.
9. Let $f(x) = |x|$ and $g(x)$=3-(x^2). Use a graphing utility to sketch a graph of $f+g$ and g/f. What are the domains and ranges of $f+g$ and g/f?
10. Let $f(x) = (x+a)/(bx+c)$ and $g(x)$=dx.a. Find f($g(x)$).
    b. What is the domain of f($g(x)$)?

2.8 Headline text

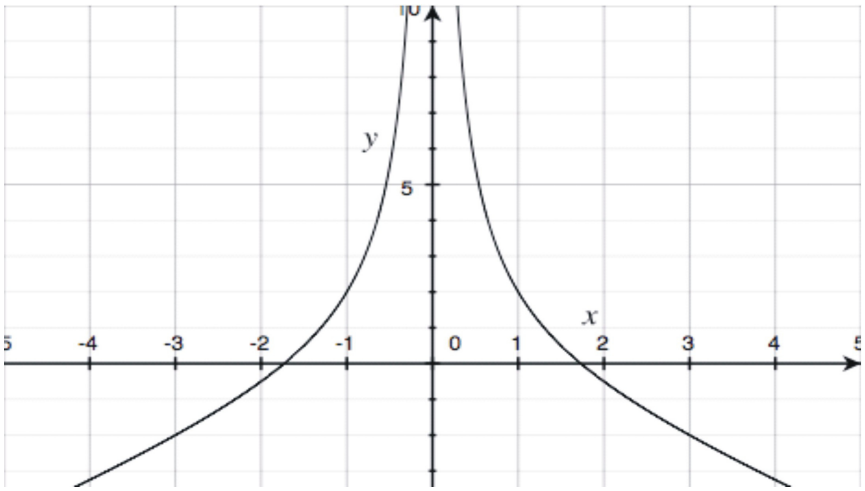
Review Answers

1. (1/2)x^2 +5x+5
2. (1/2)x^2-3x+5
3. 2x^3+ (3/2)x^2-5x
4. $h(x) = \frac{4x-5}{\frac{1}{2}x^2+x} = \frac{8x-10}{x^2+2x}$. The domain is the set of all real numbers, except x 0,x -2.
5. $f(g(x))$=$f(x+8)$=$(x+8)^3$$g(f(x))$=$g(x^3)$=$x^3$+8
   They are never equal.
6. The graph of $f(g(x))$ is the graph of $f(x)$, shifted 8 units to the left.
7. 20x-59
8. $g(x)$=3x+1,$h(x)$=x^4,$t(x)$=$f(g(x))$

9. The domain of $f+g$ is $\mathbb{R}$. The



range is y 3.25.



(g/f) has the domain of all real numbers except x=0. The range is all real numbers.

10. $f(g(x))= (dx+a)/(bdx+c)$.The domain of the function is all real numbers except $x = (-c)/bd$.

2.9 Functions and Mathematical Models

Learning objectives

- Categorize some realistic situations in terms of function families.
- Write a function to represent a situation.
- Use a function to determine key aspects of a situation.

Introduction

Throughout this chapter we have examined different kinds of functions and their behavior, and we have used functions to represent realistic situations. When we use a function to help us understand phenomena such as how to maximize the volume of a container or to minimize its surface area, we are engaging in mathematical modeling. In reality, scientists and social scientists use mathematical models to understand a wide variety of quantifiable phenomena, from the workings of subatomic particles, to how people will function in the economy.

In this lesson we will revisit some of the examples we have seen in previous lessons, in an effort to categorize models according to function families. We will look at several examples of models in depth, specifically in terms of how we can use a graphing calculator to help us analyze models.

Linear models

The very first example of a function in this chapter was a linear model. The equation $y=3x$ was used to represent how much money you would bring in if you sold x boxes of cookies for \$3 per box. Many situations can be modeled with linear functions. The key idea is that some quantity in the situation has a constant rate of change. In the cookie-selling example, every box costs \$3.00. Therefore the profits increase at a constant rate.

The cookie-selling model is an equation of the form $y=mx$. The function necessarily contains the point (0,0) : if we don't sell any cookies, we don't bring in any money!

Other models will be of the form $y=mx+b$. The constant b is the y-intercept of the function, and represents the value of the function when x is zero. For example, consider a situation in which you plan to save money at a constant rate of \$20 per week. If you begin to save money after receiving a gift of \$100, you can express the amount you have saved as a function of time: $S(t)=20t+100$, where t represents the number of weeks you have been saving. The function is linear because of the constant rate of change, that is, the constant savings of \$20 per week.

Notice that in both of these examples we will only consider these functions for x values 0. In the first example, x represents the numbers of boxes of cookies, which cannot be negative. In the second example, x represents the number of weeks you have been saving money. In theory we could extend this situation back in time, but the given information does not indicate that the model would make sense. This is the case because you received \$100 as a gift at a particular point in time. You didn't save that \$100 at \$20 per week.

Both of these examples also are linear functions with positive slope. In both situations, the function increases at a constant, or steady rate. We could also use a linear function to model a situation of constant decrease.

In sum, linear functions are used to model a situation of constant change, either increase or decrease. Next we will consider functions that can be used to model other kinds of situations.

Quadratic models

In lesson 2 we saw several quadratic functions that were used as models. For example, the function $A(x) = 50x - \frac{x^2}{2}$ was used to represent the area of a rectangular plot of land enclosed on three sides by 100 feet of fence. We used this model to identify the maximum possible area for the plot of land.

Because the graph is a parabola and the coefficient of x^2 is negative, it represents an x-axis reflection (i.e., it is "upside-down"). Thus we know that the vertex will be a global maximum. This maximum point represents the maximum possible area for the enclosure. In the homework problems for lesson 2, you were asked to represent other situations using quadratic models. These situations were all related to the area and perimeter of rectangles. However, other kinds of phenomena can also be modeled with quadratic functions.

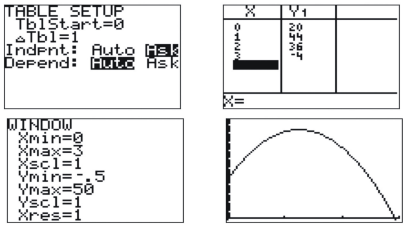
Consider for example a situation in which a ball is tossed into the air. The ball will travel up, and then it will travel down until it hits the ground. How high will the ball go? When will it reach the ground? This kind of situation can be modeled by a function of the form $h(t)=-16t^2+v_0t+h_0$. The variable t represents the time since the ball was thrown. The coefficient v_0 represents the initial velocity of the ball, and the

constant h_0 represents the initial height of the ball. The constant -16 come from the force of gravity which pulls the ball down, which is why it is negative. The following example shows a specific situation of this form.

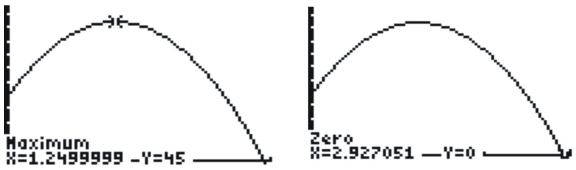
Example 1: You are standing on the roof of a building that is 20 feet above the ground. You toss a ball into the air with an initial velocity of 40 ft/sec, so that it will land on the ground, not on the roof. How high will the ball go, and when will it reach its maximum height? When will the ball hit the ground?

Solution: First we need to write a function to model the situation. Using the general form of the equation given above, we can write the function $h(t)$=-16t^2+40t+20 , where $h(t)$ represents the height above the ground..

To answer the first question, we need to examine the graph of the function. If you graph this equation on your calculator, you will need to determine a good viewing window. One way to start to determine a good window is to take into account the y-intercept of the function. In this case, the y-intercept is (0, 20). Also think about what kind of function this is: a parabola, facing downwards. This fact should lead you to think that we need to look at y-values *well above 20*. As noted earlier in the chapter, it is often useful to look at a table of values. Using the "ask" capability of the table, if you input x values of 1, 2, and 3, you will see that the function goes up to 44 at $x = 1$. The maximum value is most likely somewhere near $x = 1$. Press WINDOW, and set xmin = 0, xmax = 3. Then set Ymin =0 (or a little less, if you want to see the y-axis). Ymax should be no less than 44, though you may want to make it larger, such as 50 or more, just to be sure that you can see the vertex. Once you have set the window, press GRAPH.



Now you should see the parabola. To identify the coordinates of the vertex, you can use the MAX function in the CALC menu, introduced earlier in the chapter. Remember that the calculator will ask you to input a left bound, a right bound, and a guess for the maximum. If you use the MAX function, you should find that the coordinates of the vertex are (1.25, 45). This means that 1.25 seconds after the ball is thrown into the air, it reaches a maximum height of 45 feet.



Now, to answer the second part of the question, we need to determine when the height of the ball is 0. Graphically, we are looking for the x-intercept of the parabola. If you return to the GRAPH screen, you should see that the x-intercept is around 3. If we want to determine the exact value, or at least a good approximation of the x-intercept, we can use the ZERO function. Press 2nd TRACE to get the CALC menu, and choose option 2, ZERO. Like the MAX function, you need to input a left bound, a right bound, and a guess, although the guess is optional – just press ENTER. (Note that the calculator works this way because it is asking you to identify which x-intercept to calculate. The parabola has two x-intercepts, and other functions may have more.) Using the ZERO function, you should find that the x-intercept is approximately 2.93. This means that the ball reaches the ground in just under 3 seconds.

Like the area situations, the falling object is represented by a quadratic function that has a maximum. A

quadratic function whose vertex is a global minimum could be used to model a situation in which some quantity decreases, and then increases. In general, polynomial functions can be used to model situations in which a quantity increases and decreases, or the rate of change changes, depending on the value of x. Next we will examine cubic functions that can be used to represent such situations.

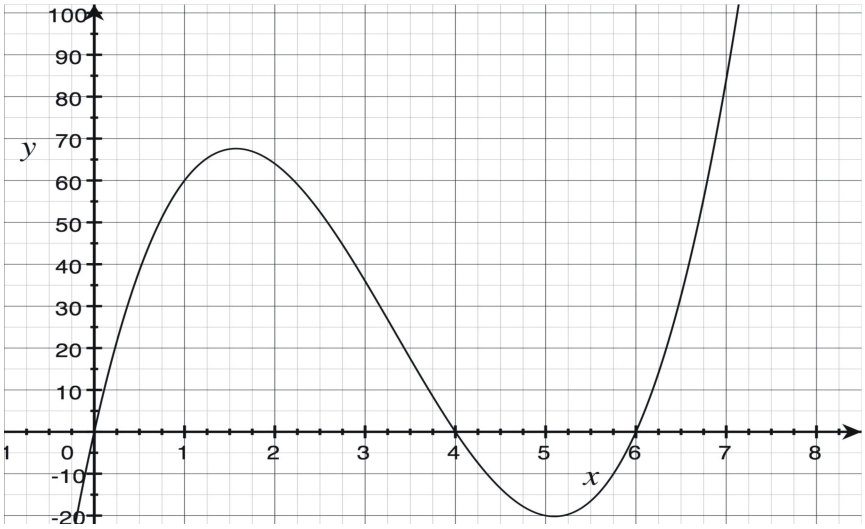
Cubic models

In lesson 3 we analyzed a situation in which a rectangular piece of cardboard was folded into a box. The folding was possible because we cut squares out of the four corners of the cardboard. We used the function $v(x)=(12-2x)(8-2x)x$ to represent the volume of the box as a function of x, the side-length of the squares cut out of the corners. If we multiply out the factors of this function we can verify that this is a cubic function:

Table 2.25:

	$v(x)=(12-2x)(8-2x)x=4x^3-40x^3+96x$

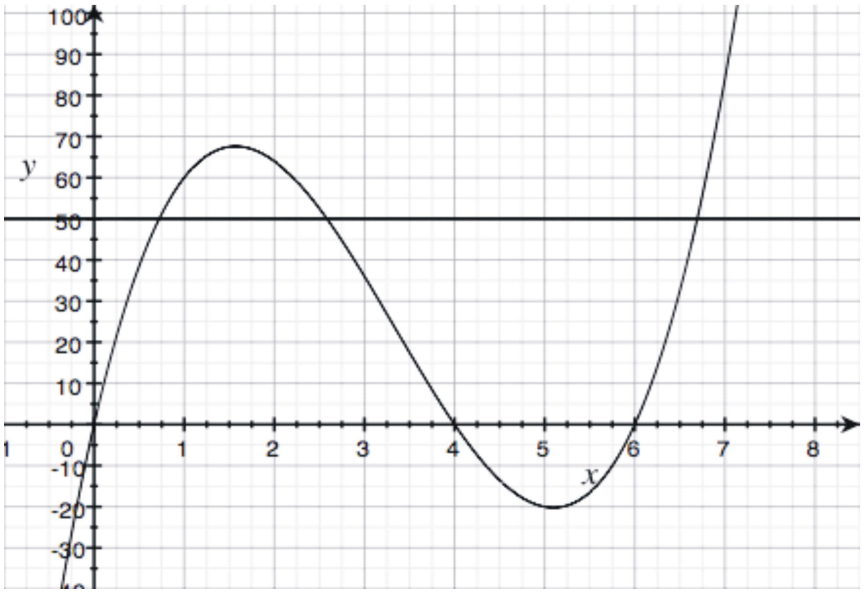
We used this function to find the maximum possible volume of the box. We can also analyze the graph to understand how the volume changes as a function of x.



When we analyzed this function in lesson 3, we only looked at the portion of the graph that looks "parabolic". We did this because the function ceases to model the situation if x is more than 4. If we cut out {4x4} squares, we would cut out one entire side of the cardboard rectangle, and we would not be able to make a box. Focusing then on the interval (0, 4), we can see that the volume of the box increases, and then decreases. If we wanted to know the volume of a box with particular dimensions, we can trace on the graph, input values into the table, or take advantage of the graph being in trace mode. That is, if you press GRAPH to view the graph, and then press TRACE, you can input x values. For example, say that you wanted to cut out squares of side-length 2.5. Press TRACE, then press 2.5, then press ENTER. At the bottom of the screen you will see $x = 2.5$ and $y = 52.5$. This tells you that the volume of the box will be 52.5in^3.

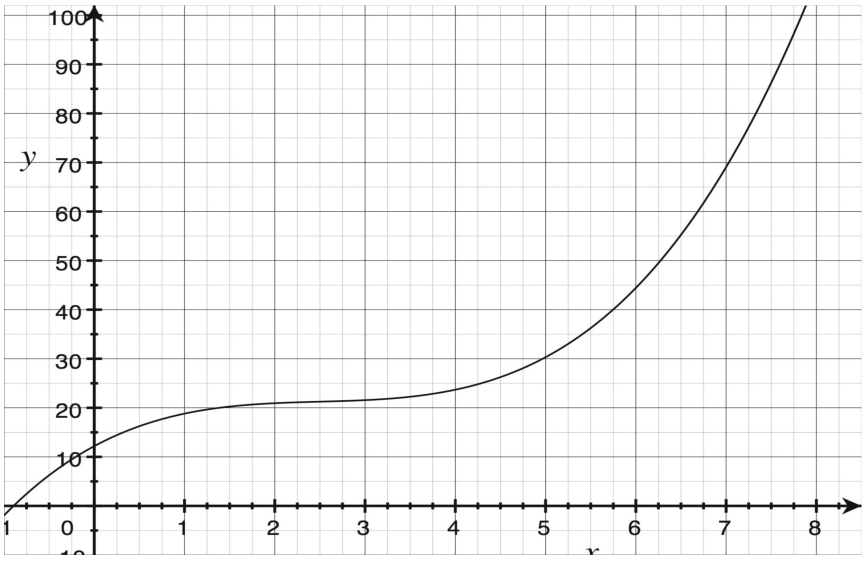
You can also determine the value of x, given a specific volume. Say that you want the box to have a volume of $50\ \text{in}^3$. One way to determine the value of x is to graph the constant function $y = 50$, and find the point where the volume function intersects $y = 50$. Press Y= and enter 50 into *Y2*. Now press GRAPH. You

should see the horizontal line $y = 50$ intersecting the volume function in several places. We are interested in the two intersection points in the interval (0, 4).



To find a good approximation of an intersection point, trace close to one of the two points. If you trace close to the first point, you will see that it is around $x = 0.8$. Now press 2^{nd}, CALC, and choose option 5, INTERSECT. The calculator will send you back to the graph screen, and ask you to choose the first curve. (The calculator does this in case you have more than two functions graphed at the same time.) You should see the cubic equation at the top of the screen. Press ENTER, and the calculator will ask you for the second curve. You should see $y = 50$ at the top of the screen. Press ENTER, and then enter a guess. (If you have already traced close to an intersection point, and you only have two functions graphed, you can simply press ENTER three times.) You should now see the coordinates of the intersection point at the bottom of the screen: x is approximately 0.723. (You can use the same process to estimate the other intersection point.)

A cubic function can be used to model situations that involve volume, but they can also be used to model situations that follow particular growth patterns. For example, a cubic function can be used the model cost of producing commercial goods. In the graph below, the x-axis represents hundreds of goods produced and the y-axis represents cost of production.



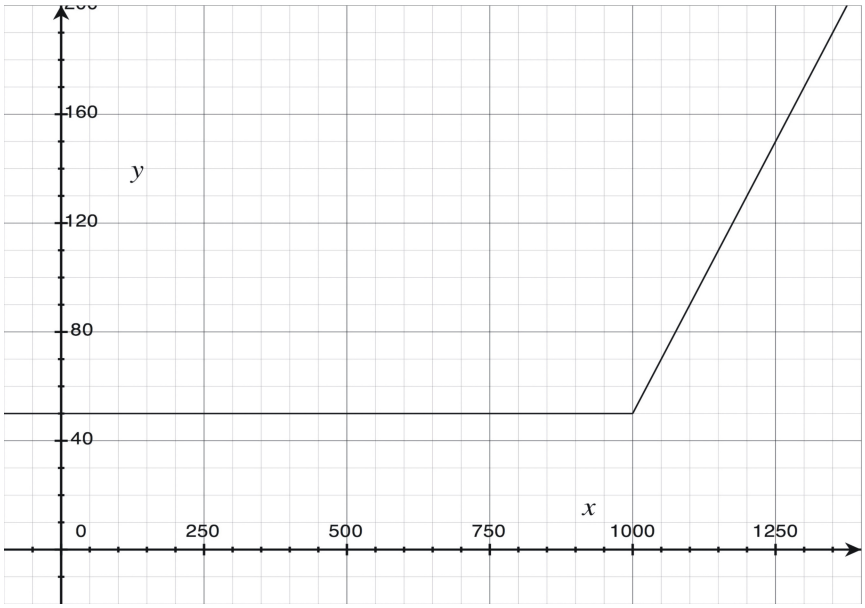
If the cost to manufacture each item was fixed, then the model would be linear. However, if we produce a larger number of items, we may be able to reduce the cost per item. For example, in the graph above, notice that the slope of the graph flattens out from about 1.5 to 4. If this graph is modeling production of x goods (in hundreds) then we can say that the manufacturing costs are increasing less quickly in this interval because we are likely spending less per item when we produce more that 150 items at a time. The costs increase faster again when we are producing more than 400 items, perhaps because we might not have the resources to produce that many items, or more people or overtime would be required. In the review questions, you will examine a specific example of a cost function.

Now we turn to other kinds of models.

Other models

Besides linear, quadratic, and cubic functions, in this chapter we have seen other functions that can also be used to model situations. For example, the equation $S(x) = (48/x) + 2x^2$ was used in lesson 2 to model the surface area of a box. In a review question in lesson 3, you wrote an equation to model the area of a rectangle inscribed under a semi-circle. This equation belonged to the square root family. In lesson 5 we examined piece-wise functions, which can be used to describe situations in which quantitative relationships are different in different intervals within the domain of the function. For example, consider a situation in which a wireless provider offers customers a monthly plan that costs \$50, but then charges \$0.40 cents per minute for every minute over 1000 included day time minutes. We can model the monthly cost, C, of the plan as a function of m, the number of day time minutes you use:

$$C(m) = \begin{cases} 50, & m \leq 1000 \\ 50 + 0.40(m - 1000), & m > 1000 \end{cases}$$



This function is comprised of a constant function, and a linear function with slope 0.40. If in a given month you use 1000 minutes or fewer, your monthly cost is a constant \$50. If you use more than 1000, each additional minute influences the value of C. For example, if you use 1,020 minutes, your cost is:

Table 2.26:

	$C(1020)$ = 50+.40(1020-1000)
	= 50+.4(20) = 50+8 =\$58.00

It is important to note that in this kind of situation, the time used may to be rounded to the nearest minute. So, for example, if you use 20.5 minutes, you will be charged for 21 minutes. You will be asked to examine this kind of function in one of the review questions.

In the previous lesson, we examined function operations, including composition of functions. We can use composition to model situations that involve more than one function. Consider for example a situation in which you are painting walls. You can express the number of gallons of paint you need for a given wall as a function of the square footage of the wall. If we assume that a gallon of paint can cover approximately 350 square feet, then the equation $g(x)$= x/350 can be used to model the number of gallons needed for a wall that is x square feet in area. Now you need to buy paint, which usually comes in gallons. If we choose a paint that costs \$25 per gallon, then the equation $c(g)$=25g represents the cost of g gallons of paint. If we compose these functions, we obtain a new function that represents the cost of painting a wall that is x square feet in area: $c(g(x))$=c (x/350) = 25x/350 = x/14 .

When developing a model, it is always a good idea to step back and ask yourself if the specific equation makes sense, and if the type of equation makes sense, given the situation. For example, notice that all three functions involved in the previous example are linear. This makes sense: they all involve a constant rate of change. If we return to the cell phone example, this model makes sense to people who have ever gone over the minutes included in their plan! Finally, if we consider the surface area equation, we are reminded that we have to carefully define the domain of the model. The graph of $S(x)$= (48/x)+2x^2 includes a portion that takes on negative x values, which don't make sense if we are modeling surface area.

Lesson Summary

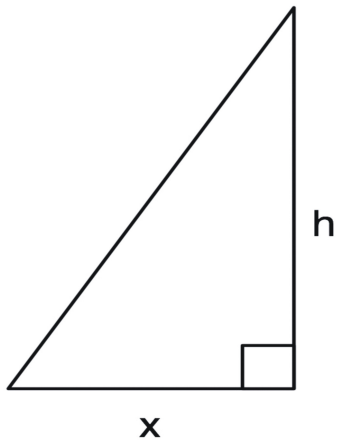
In this lesson we have revisited examples of mathematical models that we have seen throughout this chapter, and we have examined new examples as well. In each example, we considered the details of a given situation, we developed an equation to represent the quantities in the situation, and we analyzed function values to answer questions about the situation. While earlier in the chapter we focused on maximizing and minimizing quantities, here we have also examined other function values. For example, in the example of the ball being tossed into the air, we determined the time at which the ball would hit the ground by calculating the x-intercept of a parabola. In general, using a graphing calculator will help you find function values, including maxima and minima. For the review questions, you are encouraged to use the techniques discussed throughout the chapter, including using the calculator to view a graph, to analyze a table of values, and to identify key values of a function.

Points to Consider

1. What aspects of different function families make them appropriate to model different phenomena?
2. How can a graphing calculator help you analyze a model?

Review Questions

1. Consider this situation: you run a business making birdhouses. You spend $600 to start your business, and it costs you $5.00 to make each birdhouse. a. Write a linear equation to represent this situation.
   b. State the domain of this function.
   c. What does the y-intercept represent? What does the slope represent?
2. Consider this situation: you borrow $500 from a relative, and you agree to pay back the debt at a rate of $15 per month. a. Write a linear model to represent this situation.
   b. Explain why this situation is linear.
   c. Graph the function you wrote in part (a) and use the graph to determine the number of months it will take to pay off the debt.
3. You are standing on a wood platform that is 10 feet off the ground. You throw a ball into the air with an initial velocity of 38 ft/s. Write an equation to model the height of the ball as a function of time. How long will it take the ball to reach its maximum height? When will it hit the ground?
4. You are creating a garden in the shape of a triangle. You want the triangle's height to be 3 feet shorter than the length of the base. You plan to build a fence along the hypotenuse.



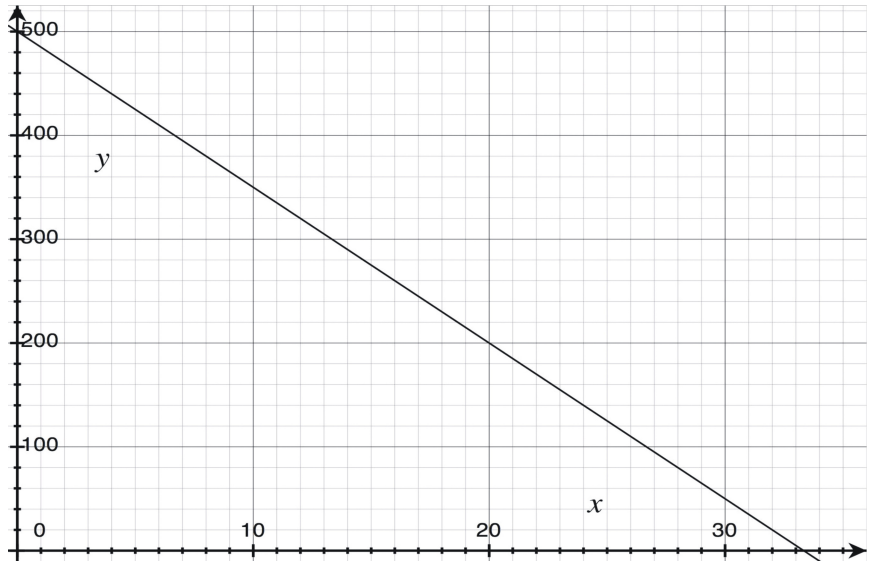
   a. Write an equation to model the length of the hypotenuse, the longest side of the triangle.
   b. What value of x will minimize the length of the hypotenuse?
5. A box is to be made by cutting squares out of the corners of a rectangular piece of cardboard. The dimensions of the cardboard are n inches by m inches. Assume that $n>m$.a. Write a model for the volume of the box.
   b. What is the largest square that can be cut out of the corners of the cardboard?
6. At your favorite gourmet store, you buy loose tea by the ounce. The store does not calculate for fractions of an ounce. Instead, they round the weight of tea UP to the nearest ounce. Your favorite tea costs $3.00 per ounce. Sketch a function to model the cost of tea as a function of the number of ounces you purchase. Explain why the function is not continuous.
7. The profits for a business can be determined by subtracting the costs from the revenue. Suppose the revenue of a business is modeled by the function $R(x) = 5x - 0.01\ x^2$, and the costs of manufacturing the product is modeled by $C(x) = 100 + 2x$, where x is the number of units of the product. a. Write a function $P(x)$ to model the company's profits.
   b. Graph $P(x)$ and determine the maximum profit.
8. Express the following situation as a composition of functions: You are running a small business making wooden jewelry boxes. It costs you $5.00 per unit to produce wooden boxes, plus an initial investment of $300 in other materials. It then costs you an additional $2.00 per box to decorate the boxes.
9. Consider the following situation: you are running a business, and you make an initial investment of $1000. It costs you $12 to manufacture each item. a. Write a function to model your total costs.

b. Write a function to model the average cost per item. Use a graphing utility to determine the minimum of this function. What does this mean?

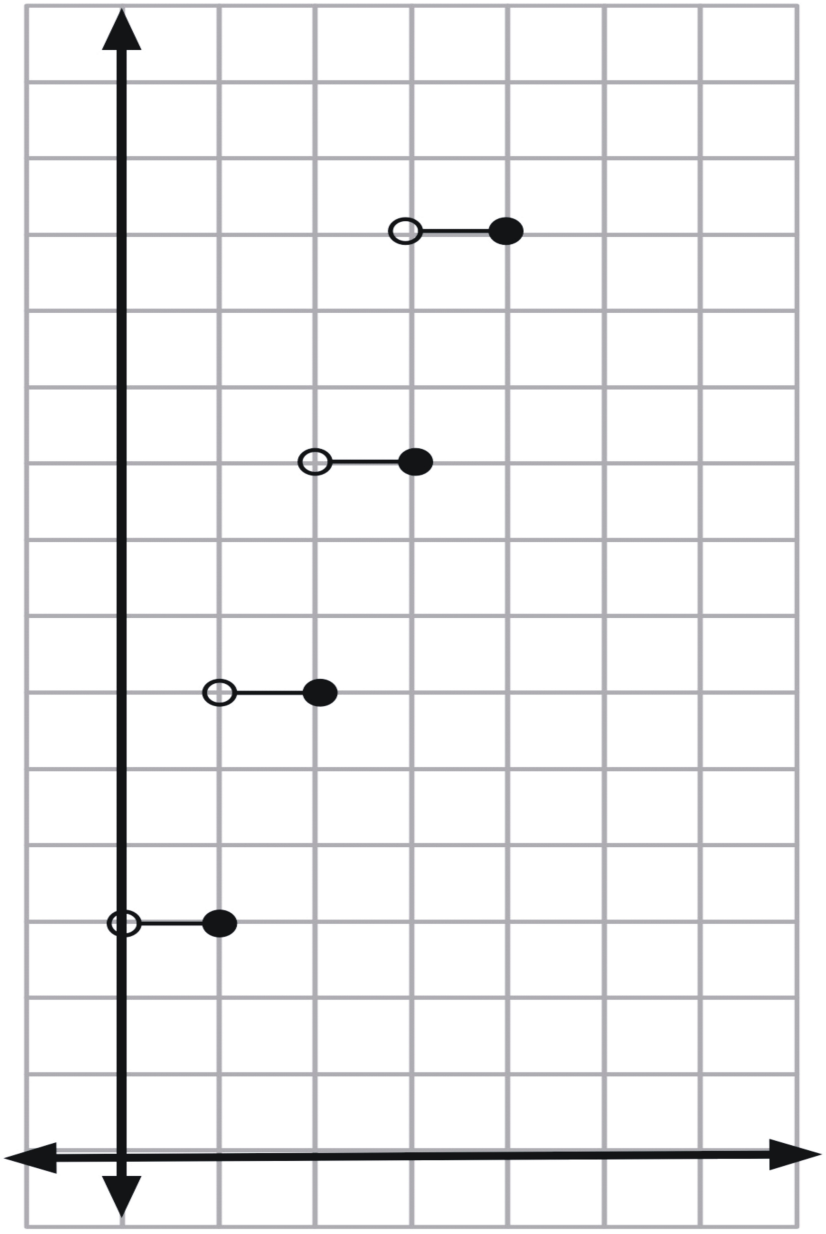
10. The function $f(x) = 0.03x^3$ - $.2x^2 + x + 4$ is used to model the costs of a running a company that manufacturers book cases. a. Use a graphing calculator to graph this function. What interval of the domain should you consider?
    b. If f represents the costs in thousands, and x represents the number of toys in thousands, what is the cost of manufacturing 2000 toys? How much is this per toy?
    c. What is the cost of manufacturing 8000 toys? How much is this per toy?
    d. If the cost is $20,000, how many units have been produced?

Review Answers

1. a. $y = 5x$+600b. D: All real numbers greater than or equal to 0.
   c. The y-intercept (0, 600) represents how much money your business has cost you before you have produced any birdhouses. The slope, 5, represents the cost per birdhouse.
2. a. $y = -15x + 500$b. The situation is linear because you are paying the debt at a constant rate.
   c.

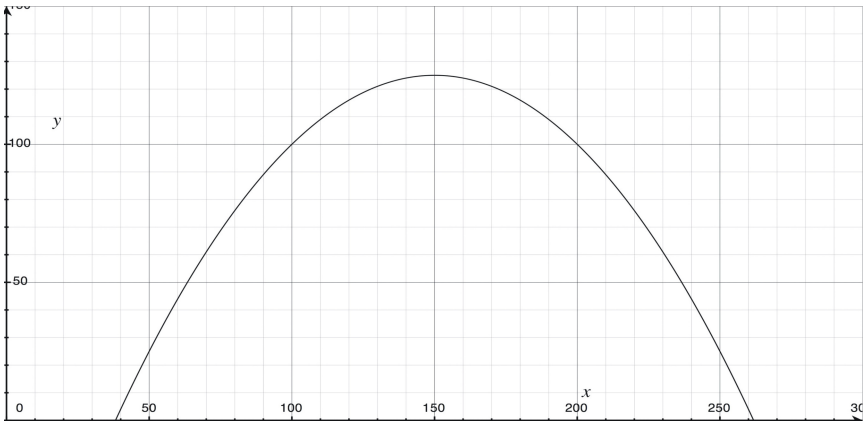


   It will take 33(1/3) months, or 34 months to pay off the debt.
3. $h(t) = -16t^2 + 38t + 10$The ball reaches maximum height around 1.2 seconds.
   The ball reaches the ground after about 2.6 seconds.
4. a. $h(x) = \sqrt{x^2 + (x-3)^2}$b. The function reaches a minimum around $x = 1.5$
5. a. $v(x) = (m - 2x)\ (n - 2x)\ (x)$.b. The side length of the square must be less than $m/2$.

6. 

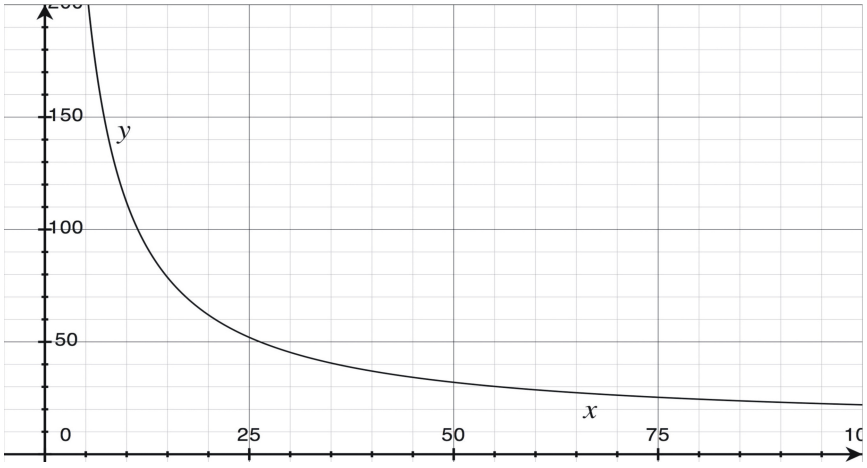
The function is not continuous because the function is constant between each whole number value of the domain, but then it "jumps up" to the next whole number value in the range. For example, 1.99 ounces of tea and 2.00 ounces both cost \$6.00, but 2.01 ounces costs \$9.00.

7. a. $P(x) = R(x) - C(x) = -0.01x^2 + 3x - 100$

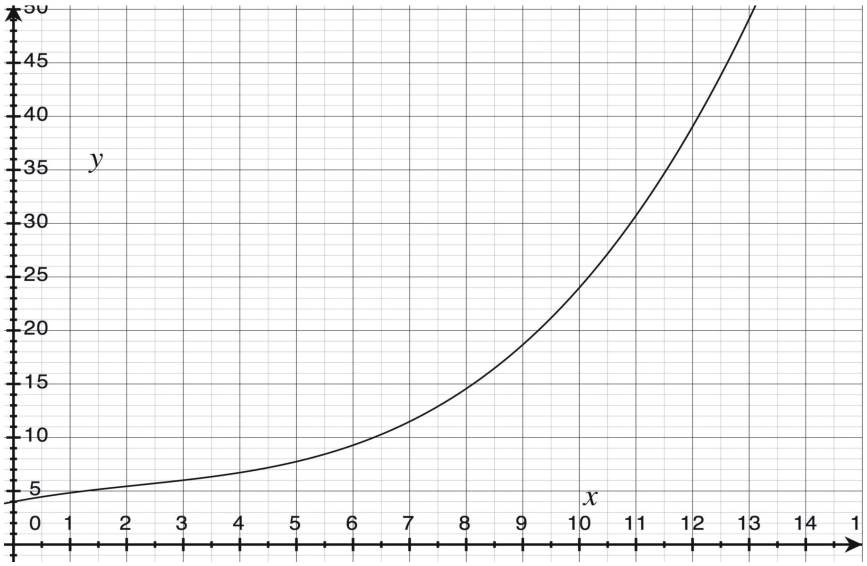


b. The maximum profit is 125 (usually in thousands, or another larger unit!)

8. Initial cost function: $C_1(x) = 5x + 300$Second cost function $C_2(x) = 2x$
   Composition: $C(x) = C_1(C_2(x)) = 5(2x) + 300 = 10x + 300$.
9. a. Cost: $C(x) = 12x + 1000$b. Average cost: $A(x) = (12x + 1000)/x$. This function does not have a minimum value. The function is asymptotic to the line $y = 12$. This means that the cost per unit will always be more than \$12.



10. a.



Consider x values 0.b. \$5,440; \$2.72

c. $14,560; $1.82
d. About 9,270 units.

Chapter 3

Analyzing Polynomial and Rational Functions

3.1 Quadratic Functions

Learning objectives

- Learn how to solve a quadratic equation by the three methods: factoring, completing the square, and the quadratic formula.
- Understand the properties of quadratic functions and how to use them to make a graph.

Standard Form of Quadratic Function

We know from physics, if you throw a ball up in the air from, say, 10 meters above ground with an initial speed of 12 meters per second, then its height h above the ground t seconds later is given by the function $h(t) = -4.9t^2 + 12t + 10$. This function is an example of a *quadratic function*.

Definition: Quadratic Function

A function f defined by $f(x) = ax^2 + bx + c$, where a, b, and c are real numbers and $a \neq 0$, is called a *quadratic function*.

The defining characteristic of a quadratic function are that it is a polynomial whose highest exponent is 2.

The form of the quadratic function above, $f(x) = ax^2 + bx + c$, is called the *standard form* of a quadratic function. There are several other ways to write quadratic functions such as in *vertex form*, $f(x) = a(x - h)^2 + k$, and in *factored form*, $f(x) = a(x - r_1)(x - r_2)$. You can move between forms of quadratic functions using algebra and you will see in this chapter that we can use each of these forms to graph a quadratic functions.

The *y−intercept* of a quadratic function in standard form is $(0, c)$ and it is found by substituting 0 for x in $f(x) = ax^2 + bx + c$.

Graph Quadratic Functions Using Transformations

The shape of the graph of any quadratic function is called a *parabola*, and it is the same as the shape of $f(x) = x^2$ (see below). However, the basic graph may be moved, reflected, or stretched depending on

the values of the coefficients a, b, and c. Examples of graphs of quadratic functions are shown in below. Notice that they all have the same shape, but some parabolas open upward and have a *minimum* point (called the *vertex* point) and some open downward and have a *maximum* vertex point. In calculus, those minimum and maximum points are usually called the *extreme* points of a function. As the name implies, at its extreme point(s), a function has a maximum or a minimum height.

Table 3.1:

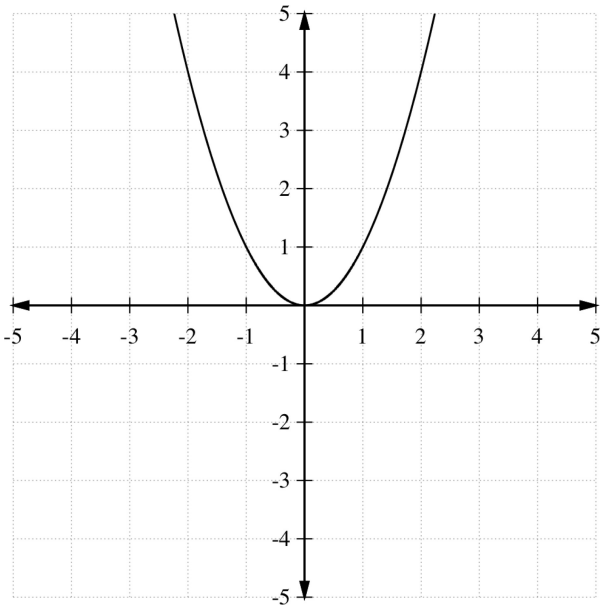
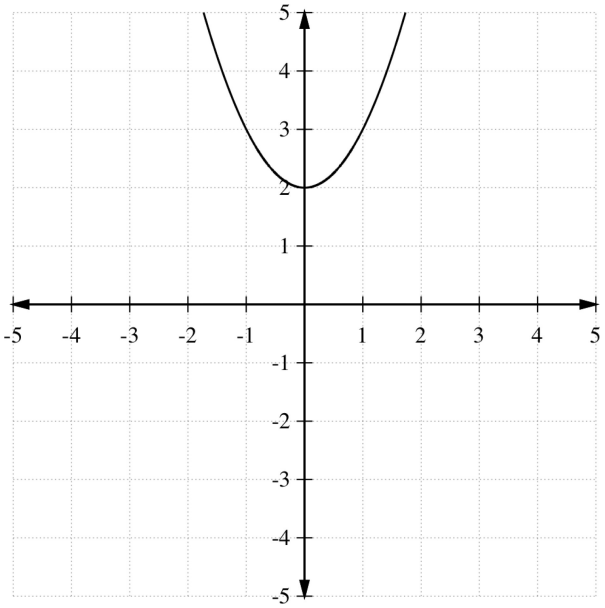
Graph	**Equation**
	$f(x) = x^2$
	$f(x) = x^2 + 2$

Table 3.1: (continued)

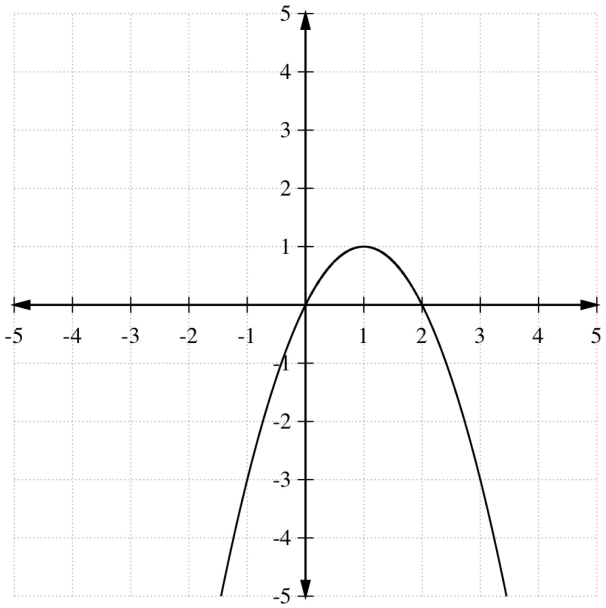
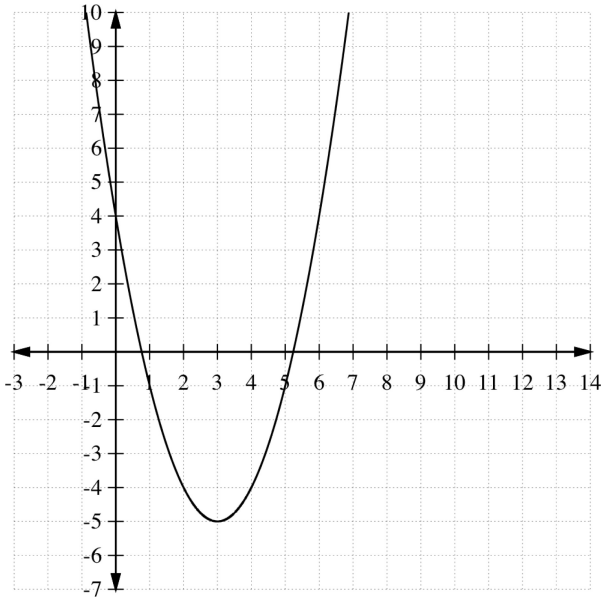
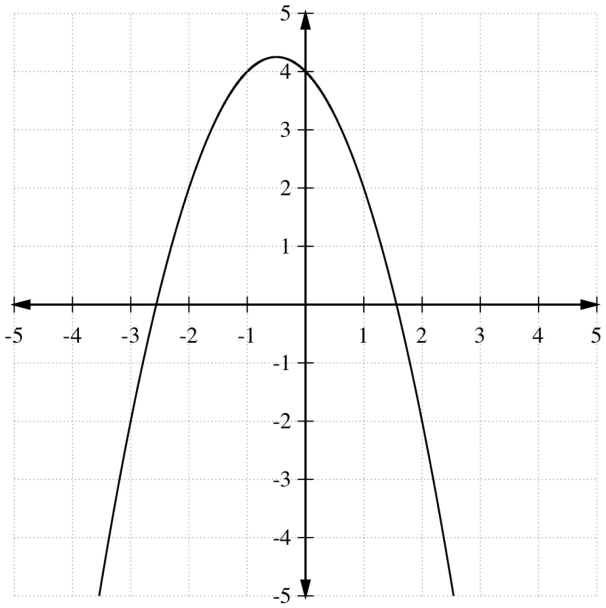
Graph	Equation
	$f(x) = -x^2 + 2x$
	$f(x) = x^2 - 6x + 4$

Table 3.1: (continued)

Graph	Equation
	$f(x) = -x^2 - x + 4$

One way to graph a quadratic function by hand is to translate the basic graph of $f(x) = x^2$ using transformations. You can apply rules about vertical shift, horizontal shift, and "stretching" of the basic parabolic shape. This is simplest to do if the quadratic function is written in vertex form by completing the square ((insert cross reference?)).

Summary of Vertex Form:

Given a quadratic function in the form $f(x) = a(x-h)^2 + k$:

- The vertex is at (h, k)
- The parabola opens up if $a > 0$
- The parabola opens down if $a < 0$
- The parabola is narrower than $y = x^2$ if $|a| > 1$
- The parabola opens wider than $y = x^2$ if $|a| < 1$

Example 1

Graph $g(x) = x^2 + 6x + 7$ using transformations

Solution:

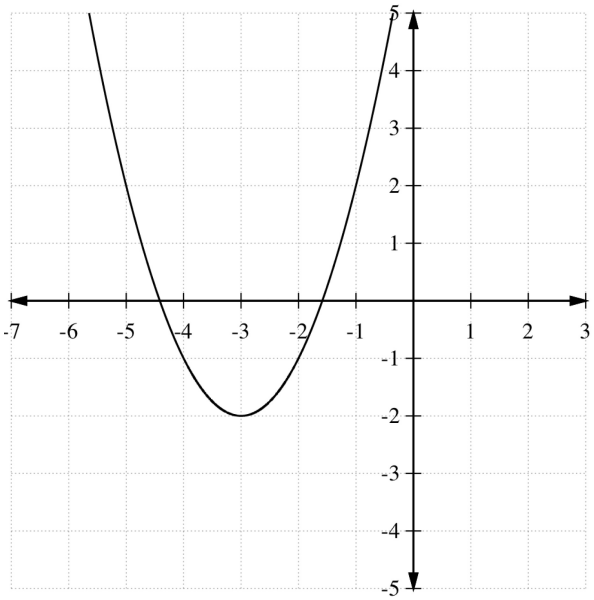
First we need to complete the square to write this function in vertex form. Add and subtract $\left(\frac{b}{2}\right)^2$ to the right hand side of the equation:

$$\begin{aligned} g(x) &= x^2 + 6x + 7 \\ &= x^2 + 6x + 9 + 7 - 9 \end{aligned}$$

Now factor the right hand side:

$$g(x) = (x+3)^2 - 2$$

Thus, $a = 1$ and the vertex of this parabola is (-3, -2). We know that the parabola opens up with the same width as $y = x^2$ and it has a minimum value at the vertex. the graph of the parabola is below.



Graph Quadratic Functions Using Vertex, Axis, Intercepts

Vertex of a parabola

Recall that the graph of a quadratic function is a parabola. In the the standard form of a quadratic function the $x-$coordinate of the vertex of the parabola is given by the equation

$$x = -\frac{b}{2a}$$

The $y-$coordinate of the vertex is found with

$$y = f\left(-\frac{b}{2a}\right)$$

Axis of Symmetry of a Parabola

A parabola has reflective symmetry about a vertical line through the vertex. The vertical line $x = -\frac{b}{2a}$ is also the parabola's *axis of* symmetry.

Example 2

Find the vertex and graph the quadratic function $g(x) = x^2 - 8x + 12$

Solution: The $x-$coordinate of the vertex is $x = -\frac{-8}{2} = 4$.

The $y-$coordinate of the vertex is $g(4) = (4)^2 - 8(4) + 12 = 16 - 32 + 12 = -4$. Thus the vertex is at (4, -4).

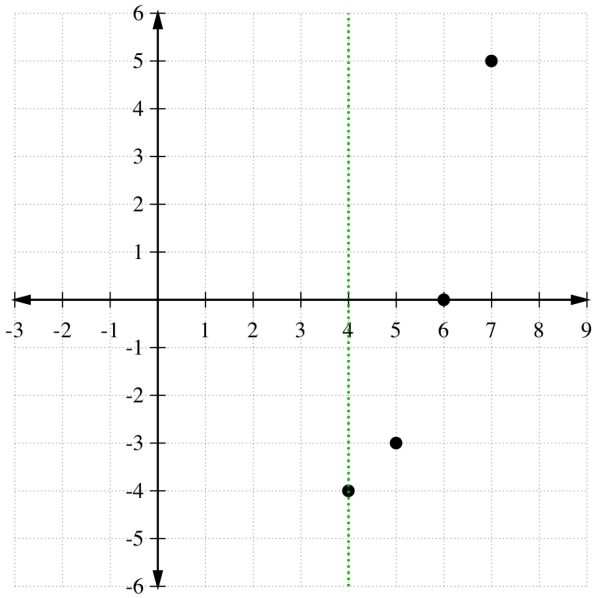
To graph the parabola, we will make a table of points staring with the $x-$coordinate of 4:

Table 3.2:

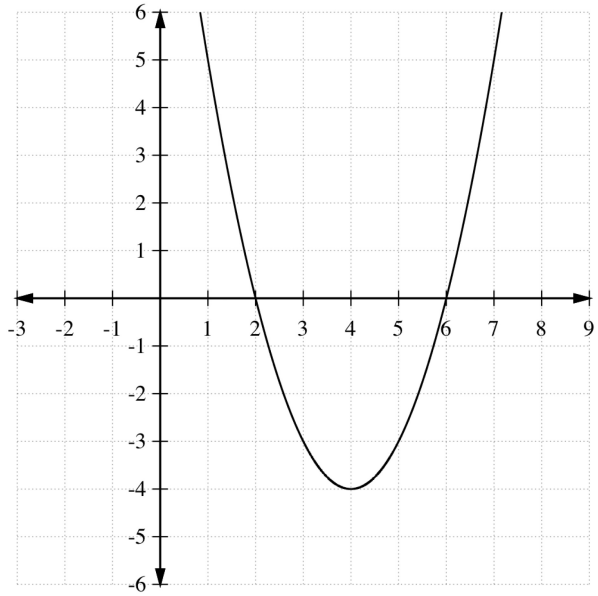
x	$y = g(x)$
4	-4

Table 3.2: (continued)

x	$y = g(x)$
5	-3
6	0
7	5



Now we can use the symmetry of $g(x)$ to fill in this table for $g(3)$. Note that $g(3) = g(5) = -3$. Likewise, $g(2) = g(6) = 0$. The final graph is below.



Intercepts of Quadratic Functions

Just like linear functions, the $y-$ and $x-$intercepts of a quadratic function can be calculated by setting $x = 0$ for the $y-$intercept and setting $y = 0$ for the $x-$intercept. For example, to locate the $y-$intercept, substitute $x = 0$ in $f(x) = ax^2 + bx + c$ to obtain $y = f(0) = c$. The $x-$intercepts are located by setting

$y = 0$ and then solving the quadratic equation $ax^2+bx+c = 0$, which can be solved by several methods that you have already studied from your previous mathematics courses: the *factoring, completing the square*, or using the *quadratic formula*. As a reminder, the quadratic formula is

Quadratic Formula

Given a quadratic function $f(x) = ax^2 + bx + c$, the x–intercepts of the function are:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

All three methods–factoring, completing the square, or using the quadratic formula–give the ***x–intercepts*** or the ***roots*** or ***the solution set*** or *the* ***zeros*** of the quadratic equation (you can tell this is an important concept since there are so many words for it!). The following example gives you a quick review of how to use those three methods.

Graphing Quadratic Functions

- The quadratic function is defined as $y = f(x) = ax^2 + bx + c$ and has an extreme point (the vertex) at the (x, y) coordinates

$$\left(\frac{-b}{2a}, f\left(\frac{-b}{2a}\right)\right)$$

If $a > 0$ (positive), then the extreme point is a minimum and the parabola opens upwards.

If $a < 0$ (negative), then the extreme point is a maximum and the parabola opens downwards.

- Find the x– and y–intercepts by setting $f(x) = 0$, and $f(0) = y$, respectively.

Example 4

Sketch the graph of the function $y = f(x) = x^2 + 2x - 3$.

Solution

Let's first find the intercepts. For the y–intercept, if $x = 0$, then $f(0) = -3$, or $y = -3$, so the y–intercept point is (0, -3).

Now, for the x–intercepts, if $y = f(x) = 0$, then $x^2 + 2x - 3 = 0$, or

$$x^2 + 2x - 3 = (x + 3)(x - 1) = 0$$

so that $x = -3$ and $x = 1$ are the x–intercepts, that is (-3, 0) and (1, 0).

The vertex (extreme point) is at

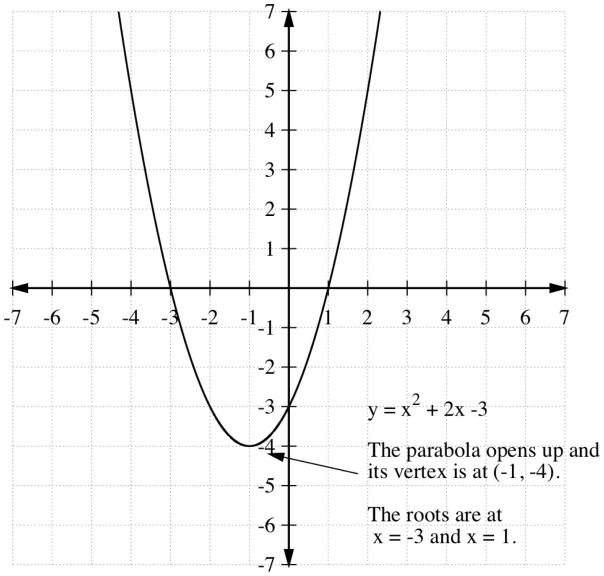
$$x = \frac{-b}{2a} = \frac{-2}{2(1)} = -1$$

Since

$$\begin{aligned} f(-1) &= (-1)^2 + 2(-1) - 3 \\ &= -4 \end{aligned}$$

The vertex is (-1, -4).

Since the coefficient of x^2 is positive, $a > 0$, the extreme point is a minimum and the parabola opens up. From this information, we can make a rough sketch of the parabola containing the points determined above. Notice that the range of the function is $y \geq -4$.



Example 5

Sketch the graph of the quadratic function $f(x) = -x^2 + 4x$.

Solution

To find the $y-$intercept, set $x = 0$, and $f(0) = -(0)^2 + 4(0) = 0$. Thus the parabola intercepts the $y-$axis at the origin. The $x-$intercept is obtained by setting $y = 0$, thus $-x^2 + 4x = 0$.

Factoring,

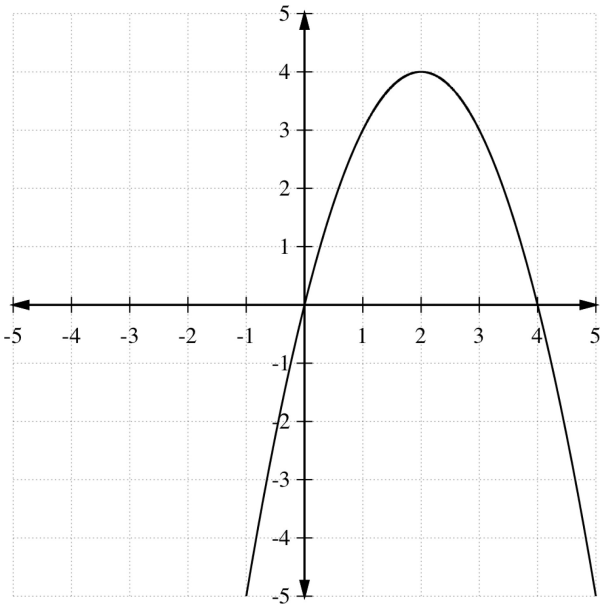
$$-x^2 + 4x = -x(x - 4) = 0$$

so that $x = 0$ and $x = 4$ are the $x-$intercepts.

We have $a = -1$ and $b = 4$, so that the extreme point occurs when

$$x = \frac{-b}{2a} = \frac{-4}{2(-1)} = 2$$

Since $f(2) = -(2)^2 + 4(2) = -4 + 8 = 4$, then $(2, 4)$ is the extreme point. It is a maximum point since $a = -1 < 0$ and the parabola opens down. Finally, the graph can be obtained by sketching a parabola through the points determined above. From graph, the range of the function is $y \leq 4$.



Example 6

A projectile is shot from the ground level. If $h(t) = 121t - 4.9t^2$, where t is the time in seconds and $h(t)$ is the height of the projectile (in meters) above the ground at time t. Find (a) h when $t = 1$ second, (b) the maximum height reached by the projectile, and (c) the graph of the trajectory of the projectile.

Solution

(a) At $t = 1$ second, we have $h(1) = 121(1) - 4.9(1)^2 = 121 - 4.9 = 116.1$ *meters*.

(b) To find the maximum height reached by the projectile, we use the formula $x = \frac{-b}{2a}$ to find the time at which the maximum height occurs (but here we use t instead of x for the independent variable):

$$t = \frac{-(121)}{2(-4.9)} = 12.3 \text{ seconds}$$

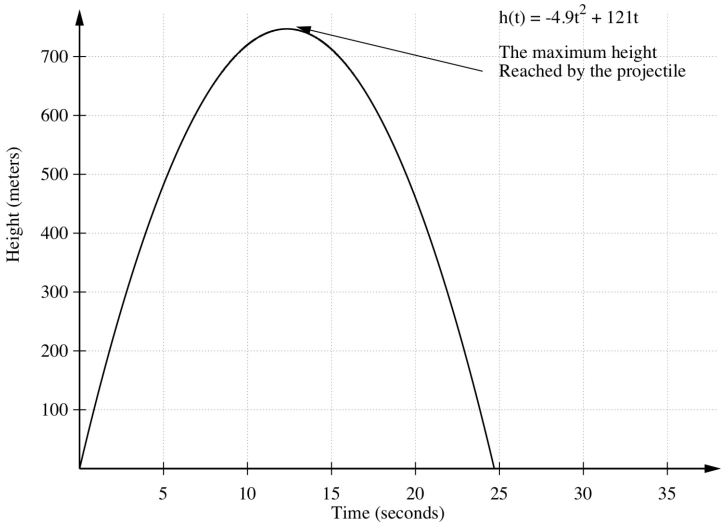
In other words, the maximum height is reached at time 12.3 seconds. That tells us when the projectile reached its maximum height. To find the actual maximum height, we simply substitute this value into the height equation

$$h(12.3) = 121(12.3) - 4.9(12.3)^2 = 747 \text{ meters}$$

(c) In order to graph it accurately, we need to find the $t-$intercepts. They can be found by setting $h = 0$:

$$121t - 4.9t^2 = 0$$
$$(121 - 4.9t)t = 0$$

So the intercepts are at $t = 0$ seconds and $t = 24.7$ seconds. The $h-$intercept occurs when $t = 0$, so $h(0) = 121(0) - 4.9(0)^2 = 0$. From this information, we can construct the graph, as shown in the Figure 5. Notice that the graph does not extend beyond the interval on the $t-$axis. (why?)



Summarize Analysis of Quadratic Functions

Example 7

Find the roots (or x–intercepts) of the quadratic equation $x^2 - 5x + 6 = 0$ by using the three methods mentioned above.

Solution

1. The Factor Method

The factor method is based on writing the quadratic equation in *factored form*. That is, as a product of two linear expressions. So our equation maybe solved by the following way:

$$x^2 - 5x + 6 = 0$$
$$(x - 3)(x - 2) = 0$$

Recall, that

$$a \cdot b = 0 \text{ if and only if } a = 0 \text{ and } b = 0$$

This tells us that

$$x - 3 = 0$$

or

$$x - 2 = 0$$

which give the roots (or *zeros*)

$$x = 3$$

and

$$x = 2$$

In other words, the solution set is $\{2, 3\}$.

2. Completing the Square Method

To solve the above equation by completing the square, first move the "c" term to the other side of the equation,

$$x^2 - 5x = -6$$

Next, make the left-hand side a "perfect square" by adding the appropriate number. To do so, take one-half of the coefficient of x (the b coefficient) and square it and then add the result to both sides of the equation:

$$\begin{aligned} b &= -5 \\ \frac{b}{2} &= \frac{-5}{2} \\ \left(\frac{b}{2}\right)^2 &= \frac{25}{4} \end{aligned}$$

Adding to both sides of the equation,

$$\begin{aligned} x^2 - 5x + \frac{25}{4} &= -6 + \frac{25}{4} \\ x^2 - 5x + \frac{25}{4} &= \frac{1}{4} \\ \left(x - \frac{5}{2}\right)^2 &= \frac{1}{4} \end{aligned}$$

this last equation can be easily solved by taking the square root of both sides,

$$\left(x - \frac{5}{2}\right) = \sqrt{\frac{1}{4}} = \pm\frac{1}{2}$$

Hence

$$x = \pm\frac{1}{2} + \frac{5}{2}$$

and the solutions are

$$x = 3$$

and

$$x = 2$$

which are identical to answers of the factor method.

3. The Quadratic Formula

The quadratic formula is the most useful method in applications for solving quadratic functions because it works for all quadratic equations. Therefore, it is highly encouraged that you memorize the formula.

Recall, if $ax^2 + bx + c = 0$ where a, b, and c are real numbers and $a \neq 0$, then the roots of the equation can be determined by the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Here the coefficients are $a = 1, b = -5$, and $c = 6$.

Substituting into the quadratic formula, we get

$$\begin{aligned} x &= \frac{-(-5) \pm \sqrt{(-5)^2 - 4(1)(6)}}{2(1)} \\ &= \frac{5 \pm \sqrt{25 - 24}}{2} \\ &= \frac{5 \pm 1}{2} \\ &= 3 \text{ or } 2 \end{aligned}$$

which is, again, identical to our two solutions above.

Applications, Technological Tools

You can graph quadratic functions using your computer's graphing program or with a TI-83/84 Calculator. Below we give basic directions for graphing a quadratic function and finding the vertex and roots ($x-$intercepts) using the functions on a TI-83 or TI-84 Graphing calculator.

Graphing

1. As with graphing any function on the calculator, you enter the function's equation by using the $Y =$ button.
2. Set an appropriate window using the **WINDOW** menu. You can use the vertex and $y-$intercept to help you find the best values for **XMAX, XMIN, YMAX, YMIN,** and the $X-$ and $Y-$scales for viewing the graph of the parabola. Alternatively, you can set a standard window and use zooming to view the graph.
3. Press **GRAPH**.
4. Adjust and refine the view using **ZOOM** or by changing the **WINDOW** settings.

((Note: Insert 3 screen shots: 1) the $Y =$ menu with a quadratic function in Y_1, 2) The **WINDOW** menu, and 3) A view of the graph.))

Finding the Vertex

You can use functions built into the calculator to find the vertex of any parabola you graph.

1. Follow the directions above to graph a quadratic function.
2. From the Graph screen, press **2ND TRACE** (This is the **CALC** Menu).
3. Scroll down and choose **MINIMUM** or **MAXIMUM** depending on whether the vertex is a minimum or maximum.
4. The calculator takes you to the graph and prompts **LEFT BOUND**? Use the arrow keys ($<$ or $>$) to place the cursor to the left of the vertex (or enter an $x-$coordinate by typing one in), and press **ENTER**
5. The calculator then prompts **RIGHT BOUND**? Use the arrows to place the cursor to the right of the vertex and press **ENTER**.

6. Finally, the calculator will prompt you for a guess with **GUESS**? You can either enter an $x-$value close to the vertex, or press **ENTER** again if your bounds were relatively good.
7. The calculator will display the $x-$ and $y-$coordinates of the vertex.

((Note: Add 2 screen shots: 1) the **CALC** menu in step #2, 2) A graph with **LEFTBOUND**? prompt (step #4), 3) A graph showing the vertex as **MAX/MIN** (i.e. after doing step 7)))

Finding the Roots

1. From the Graph screen, press **2ND TRACE** (This is the **CALC** Menu).
2. Choose **ZERO**.
3. The calculator takes you to the graph and prompts **LEFT BOUND**? Use the arrow keys ($<$ or $>$) to place the cursor to the left of the zero (or enter an $x-$coordinate by typing one in), and press **ENTER**.
4. The calculator then prompts **RIGHT BOUND**? Use the arrows to place the cursor to the right of the zero and press **ENTER**.
5. Finally, the calculator will prompt you for a guess with **GUESS**? You can either enter an $x-$value close to the zero, or press **ENTER** again if your bounds were relatively good.
6. The calculator will display the $x-$coordinate of the zero (root).

Note: Repeat steps 1-6 to find the other root of the quadratic.

((Note: Add 3 screen shots showing the calculator after Step 2, Step 4, and Step 6.))

Exercises

1. Solve each equation by factoring.
   (a) $x^2 = 64$
   (b) $4x^2 + 7x = 2$
   (c) $4x^2 - 17x = -4$
   (d) $x^2 + 1 = -\frac{13}{6}x$
2. Solve each equation by completing the square
   (a) $x^2 - 4x + 1 = 0$
   (b) $\frac{2}{3}x^2 - x + \frac{1}{3} = 0$
   (c) $x^2 + 2.8 = 4.7x$
   (d) $2x^2 - 3x - \frac{23}{8} = 0$
3. Solve each equation by using the quadratic formula
   (a) $0.4x^2 + x - 0.3 = 0$
   (b) $25x^2 + 80x + 61 = 0$
   (c) $(z + 6)^2 + 2z = 0$
   (d) $\left(\frac{5}{7}w - 14\right)^2 = 8w$
4. Sketch the graph of the quadratic function $f(x) = -x^2 - 4x - 4$ by locating the $x-$ and $y-$intercepts and the vertex point. Use the graph to determine the range of the parabola.
5. Sketch the graph of the quadratic function $f(x) = 2x^2 + 4x - 3$ by locating the $x-$ and $y-$intercepts and the vertex point. Use the graph to determine the range of the parabola.
6. Sketch the graph of the quadratic function $f(x) = x^2 + x + 5$ by locating the $x-$ and $y-$intercepts and the vertex point. Use the graph to determine the range of the parabola.

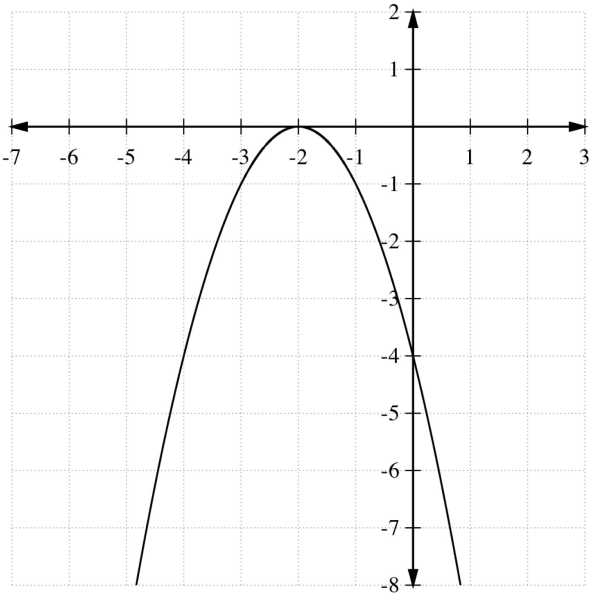
7. A projectile is fired from a cliff. If the height, h of the projectile after t seconds is given by $h(t) = -16t^2 + 96t + 256$, find

   (a) h when $t = 0$.
   (b) The maximum height reached by the projectile.
   (c) Graph $h(t)$ and show its range and domain.

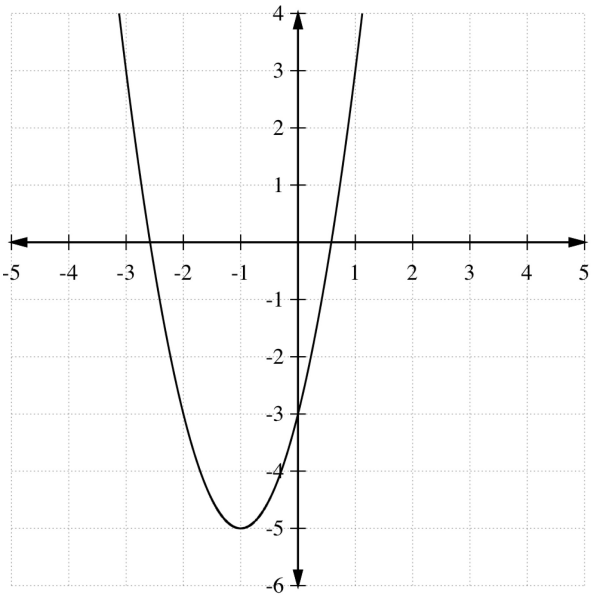
8. A surface-to-surface missile is fired and follows a parabolic path. Its path as a function of time is described by the function $p(t) = -t^2 + t + 1$. At what time the missile reaches the highest point and when does it hit the ground?
9. Use a graphing calculator to find the vertex and roots of $q(x) = -0.035x^2 - 7.25x - 12.3$. Round your answer to three decimal places.

Answers

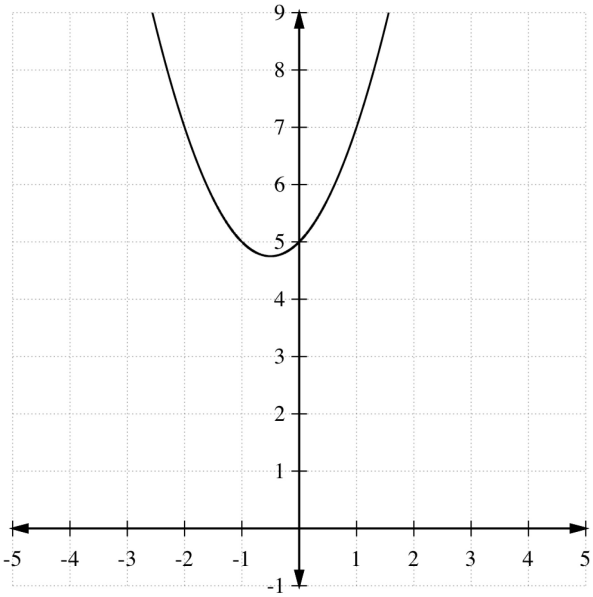
1. (a) ± 8
   (b) $x = \frac{1}{4}, -2$
   (c) $x = \frac{1}{4}, 4$
   (d) $x = \frac{-2}{3}, \frac{-3}{2}$
2. (a) $2 \pm \sqrt{3}$
   (b) $x = \frac{1}{2}, 1$
   (c) $x = \frac{7}{10}, 4$
   (d) $\frac{3}{4} \pm \sqrt{2}$
3. (a) $x = 0.275, -2.78$
   (b) $-\frac{8}{5} \pm \frac{\sqrt{3}}{5}$
   (c) $-7 \pm \sqrt{13}$
   (d) $\frac{686 \pm 196\sqrt{6}}{25}$
4. The vertex and the $x-$intercept is (-2, 0). The $y-$intercept is (0, -4).



5. Vertex: (-1, -5). $x-$intercepts: $\left(\frac{-2+\sqrt{10}}{2}, 0\right)$ and $\left(\frac{-2-\sqrt{10}}{2}, 0\right)$. $y-$intercept: (0, -3).



6. Vertex: $\left(\frac{-1}{2}, \frac{19}{4}\right)$. No x–intercepts. y–intercept: (0, 5)



7. (a) 256 ft
   (b) 400 ft
   (c)

8. At $\frac{1}{2}$ second; after 1 second
9. Vertex: (-103.571, 363.146); roots at (-205.432, 0) and (-1.711, 0).

3.2 Polynomial Functions

Learning Objectives

- Understand and apply methods to find the zeros of a factored polynomial
- Use the zeros of polynomial functions to sketch a graph of the function

Standard Form of Polynomial Function

You have already studied many different kinds of functions, for example linear functions, constant functions, and quadratic functions. All three these functions belong to a larger group of functions called the *polynomial functions*.

Polynomial Functions

If $P(x)$ is a polynomial function, then it is given by

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_1 x + a_0$$

where the coefficients $a_n, a_{n-1}, \cdots, a_1, a_0$ are real numbers and the exponents are positive integers.

We call the first nonzero coefficient a_n the *leading coefficient*. The term $a_n x^n$ is called the *leading term*. The *degree* of the polynomial is n. For example, the quadratic equation $f(x) = -2x^2 + 3x - 5$ has a leading coefficient of -2, a leading term of $-2x^2$ and of degree $n = 2$. On the other hand, the polynomial $f(x) = 1$ is a polynomial of with a leading coefficient of 1, and the leading term is $1x^0 = 1$, so the degree is $n = 0$.

A very interesting property of polynomials is that they are all *continuous*, that is, they have no holes, breaks. Informally, we say you can draw the graph of a continuous function without lifting your pencil from the paper. In calculus you will study function continuity in more depth and you will use the continuity of polynomials to simplify the analysis of other, more complicated functions.

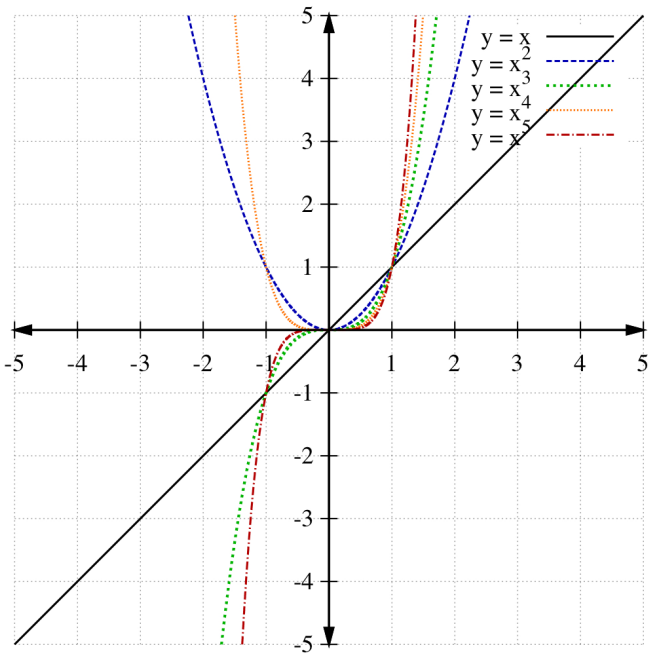
In addition to their continuity, the domain of all polynomial functions is the set of all real numbers. That is, the domain is $x \in (-\infty, \infty)$.

Sometimes it is helpful to consider a counterexample. There are many functions that we use in mathematics which are not polynomials. For example the floor function, $y = int(x)$, is a function that gives the greatest integer less than or equal to x, so $int(5.67) = 5$. This function is NOT a polynomial and it is NOT continuous.

Power Function (even, odd)

The most simple polynomial is called a power function. A *power function* is a polynomial of the form $f(x) = ax^n$ where a is a real number and n is an integer with $n \geq 1$.

If n is even, then the power function is also called "even," and if n is odd, then the power function is "odd." The graphs of the first five power functions are shown below.



Notice that each power function has only one $x-$ and $y-$intercept, the origin $(0, 0)$.

The *end behavior* of a function describes the $y-$values as x gets very large ($x \to \infty$ in symbols) or as x gets very small ($x \to -\infty$).

- For even powers n, the power function $f(x) = ax^n$ is U-shaped (like a parabola) and as $x \to \infty, f(x) \to \infty$. Likewise as $x \to -\infty, f(x) \to \infty$.
- For odd powers n, the power function goes from the third quadrant to the first quadrant (like the line $y = x$). As $x \to \infty, f(x) \to \infty$, and as $x \to -\infty, f(x) \to -\infty$.

As with quadratics and polynomials, the leading coefficient a changes the vertical "stretching" of power functions.

Graph Polynomial Functions Using Transformations

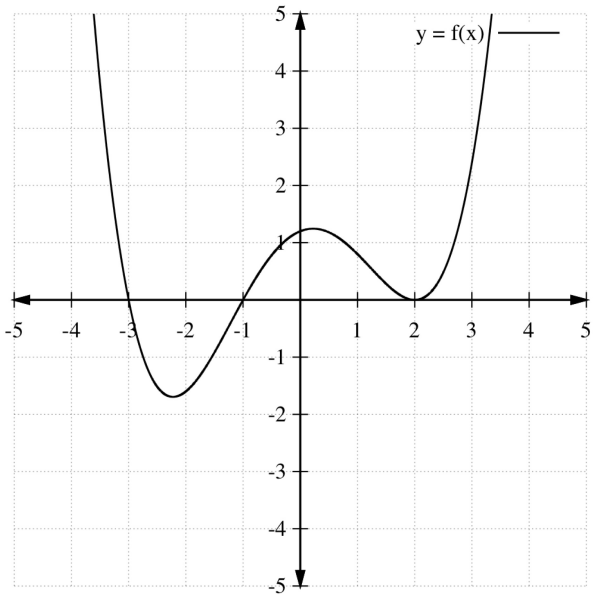
Just like quadratics, polynomial functions can be graphed using transformations of a known graph. The basic transformations are vertical and horizontal shifts and reflections about the $x-$ and $y-$axis.

Given a polynomial $p(x)$ and constant real numbers c and a

- $p(x) + c$ is a *vertical shift* of the graph of $p(x)$ by c units *up* (so the function shifts down if $c < 0$).
- $p(x - c)$ is a *horizontal* shift of the graph of $p(x)$ by c units to the *right*. (So the function shifts left if $c < 0$).
- $-p(x)$ is a reflection of the graph of $p(x)$ about the $x-$axis.
- $p(-x)$ is a reflection of the graph of $p(x)$ about the $y-$axis.
- $ap(x)$ is a vertical stretch by a multiple of a.
- $p(ax)$ is a horizontal compression by a multiple of a.

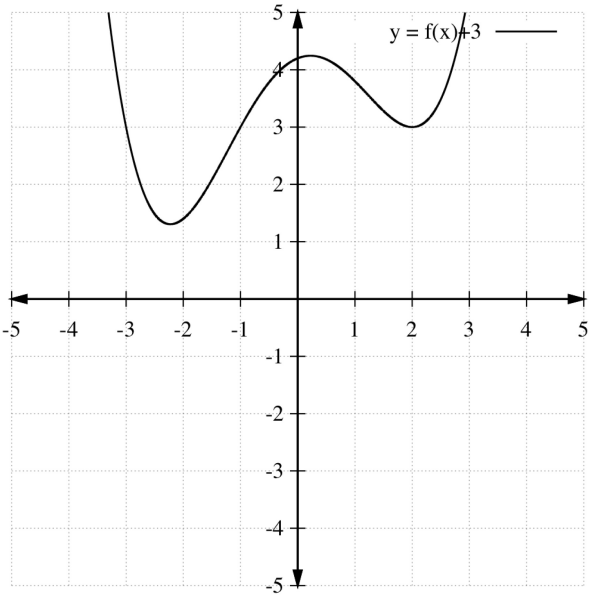
Example 1

The graph of $f(x)$ is shown below. Use the graph of $f(x)$ to graph each of the following: 1) $f(x) + 3$, 2) $f(x + 4)$, and 3) $f(-x) + 3$

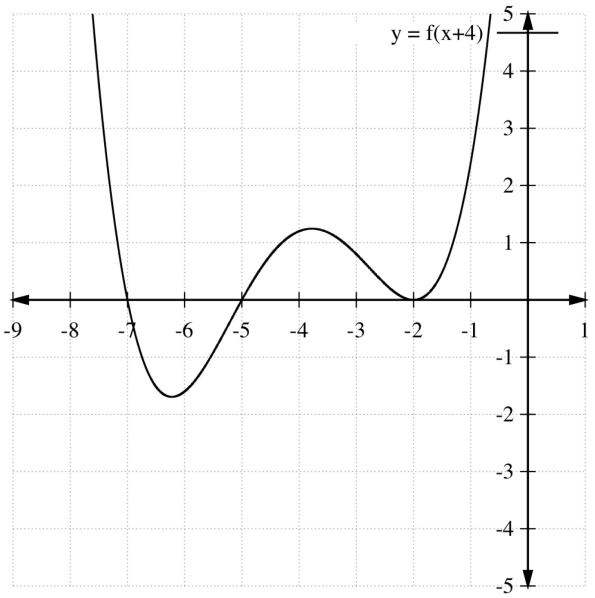


Solution

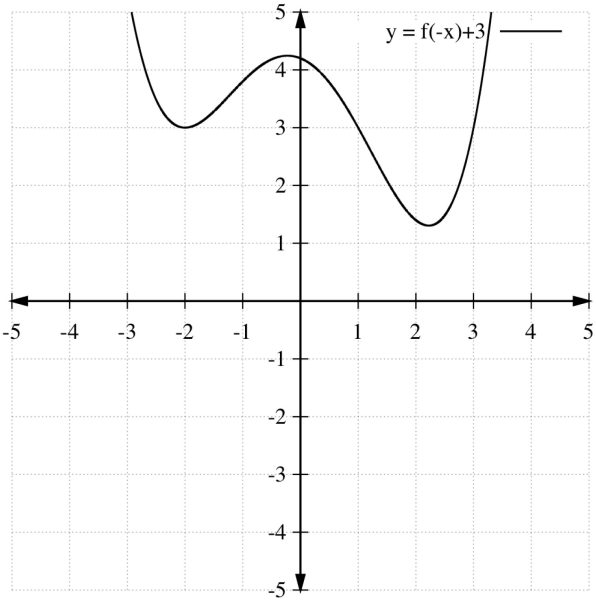
1) This is a vertical shift of $f(x)$ up by 3 units.



2) This is a horizontal shift of $f(x)$ to the left by 4 units.



3) This is a reflection of $f(x)$ about the $y-$axis and a vertical shift up by 3 units.



Graph Polynomial Functions Given Zeros ($x-$intercepts)

Recall that the zeros or the roots of a polynomial are the $x-$intercepts or the solution set of the polynomial function. For example, the polynomial

$$h(x) = x^3 + 2x^2 - 5x - 6$$

can be factored into

$$y = h(x) = x^3 + 2x^2 - 5x - 6 = (x + 1)(x - 2)(x + 3)$$

To find the zeros, we set

$$h(x) = y = 0$$

and solve for x.

$$(x+1)(x-2)(x+3) = 0$$

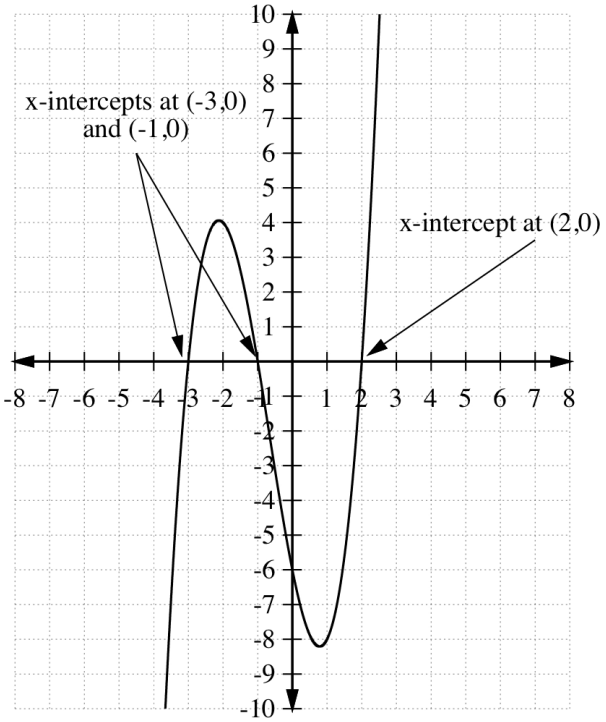
This gives

$$x+1=0$$
$$x-2=0$$
$$x+3=0$$

or

$$x=-1$$
$$x=2$$
$$x=-3$$

So we say that the solution set is $\{-3,-1,2\}$. They are the zeros of the function $h(x)$. The zeros of $h(x)$ are the $x-$intercepts of the graph $y=h(x)$ below.



The previous example illustrates a strategy for graphing a polynomial if you know the zeros of the function. If you know the zeros of a polynomial you can use "test values" between the zeros to see whether the function is above or below the $x-$aixs in the interval between the zeros. While this cannot tell you the hieght ($y-$values) of the function between the zeros, you can use the zeros to sketch an approximation of the graph. The following examples illustrate first finding the zeros, and then a method for graphing a polynomial once you know the zeros.

Example 2

Find the zeros of $g(x) = -(x-2)(x-2)(x+1)(x+5)(x+5)(x+5)$.

Solution

The polynomial can be written as

$$g(x) = -(x-2)^2(x+1)(x+5)^3$$

To solve the equation, we simply set it equal to zero

$$-(x-2)^2(x+1)(x+5)^3 = 0$$

this gives

$$x-2=0$$
$$x+1=0$$
$$x+5=0$$

or

$$x=2$$
$$x=-1$$
$$x=-5$$

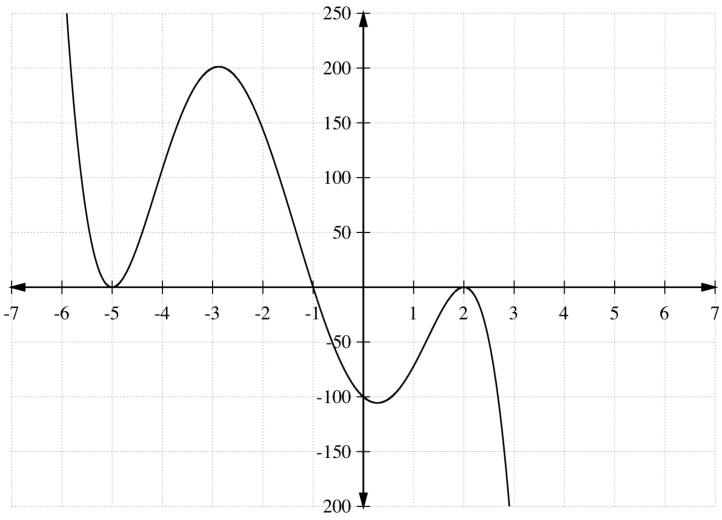
Notice the occurrence of the zeros in the function. The factor $(x-2)$ occurred twice (because it was squared), the factor $(x+1)$ occurred once and the factor $(x+5)$ occurred three times. We say that the zero we obtain from the factor $(x-2)$ has a *multiplicity* $k=2$ and the factor $(x+5)$ has a *multiplicity* $k=3$.

To graph $g(x)$, use the zeros to create a table of intervals and see whether the function is above or below the x–axis in each interval:

Table 3.3:

Interval	**Test value** x	$g(x)$	**Sign of** $g(x)$	**Location of graph relative to** x**–axis**
$(-\infty,-5)$	-6	320	+	Above
$x=-5$	-5	0	NA	
(-5, -1)	-2	144	+	Above
$x=-1$	-1	0	NA	
(-1, 2)	0	-100	-	Below
$x=2$	2		NA	
$(2,\infty)$	3	-256	-	Below

Finally, use this information and the test points to sketch a graph of $g(x)$.



Example 3

Find the zeros and sketch a graph of the polynomial

$$f(x) = x^4 - x^2 - 56$$

Solution

This is a factorable equation,

$$\begin{aligned} f(x) &= x^4 - x^2 - 56 \\ &= (x^2 - 8)(x^2 + 7) \end{aligned}$$

Setting $f(x) = 0$,

$$(x^2 - 8)(x^2 + 7) = 0$$

the first term gives

$$\begin{aligned} x^2 - 8 &= 0 \\ x^2 &= 8 \\ x &= \pm\sqrt{8} \\ &= \pm 2\sqrt{2} \end{aligned}$$

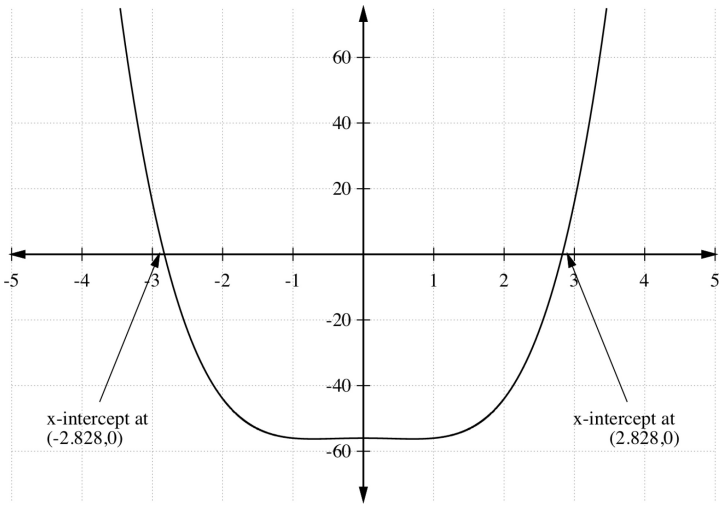
and the second term gives

$$\begin{aligned} x^2 + 7 &= 0 \\ x^2 &= -7 \\ x &= \pm\sqrt{-7} \\ &= \pm i\sqrt{7} \end{aligned}$$

So the solutions are $\pm 2\sqrt{2}$ and $\pm i\sqrt{7}$, a total of four zeros of $f(x)$.

Keep in mind that only the *real* zeros of a function correspond to the $x-$intercept of its graph. For our case in this example, only the two zeros $\pm 2\sqrt{2}$ correspond to actual $x-$intercepts (Figure 9) but $+i\sqrt{7}$ and $-i\sqrt{7}$

do not, since they are complex. These are given more attention later in the book ((Add cross-referencing link?))



Summarize Analysis and Graphing of Polynomial Functions

It is helpful to consider the following facts when graphing any polynomial function.

Two Basic Facts About the Graphs of Polynomial Functions

If $f(x)$ is a polynomial function with degree $n \geq 1$, then

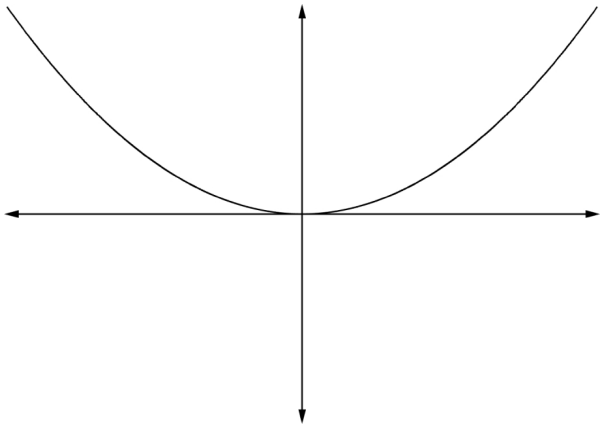
- The maximum number of real zeros ($x-$intercepts) is n.
- The maximum number of turning points is $n - 1$.

It is also helpful to learn about the behavior of the function as x becomes very large ($x \to +\infty$) or very small ($x \to -\infty$).

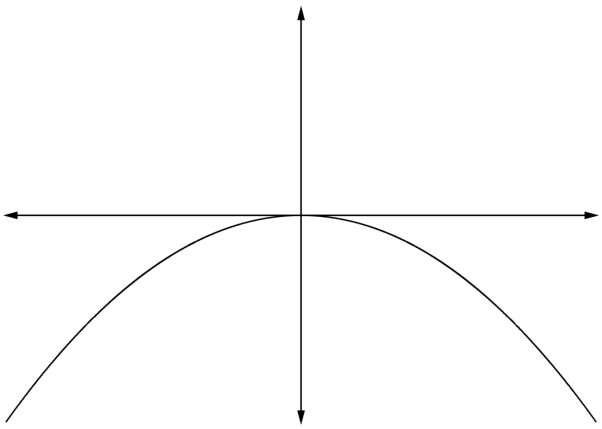
The Leading-Term Test

If $a_n x^n$ is the leading term of a polynomial. Then the behavior of the graph as $x \to \infty$ or $x \to -\infty$ can be known by one the four following behaviors:

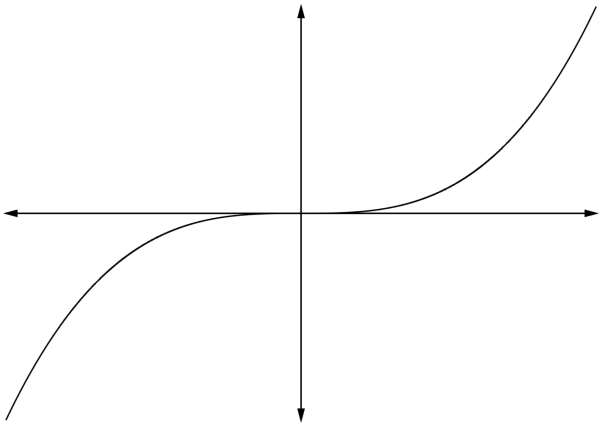
1. If $a_n > 0$ and n even



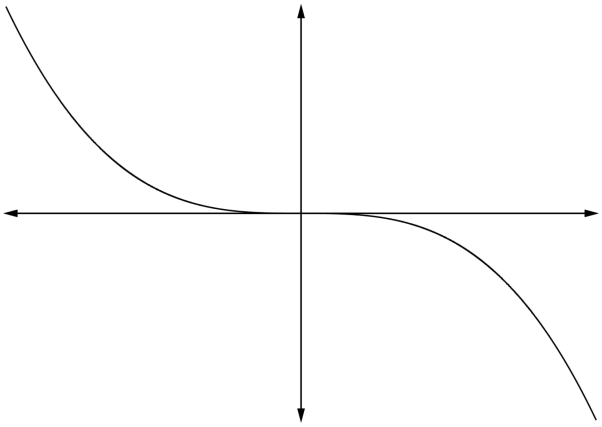
2. If $a_n < 0$ and n even



3. If $a_n > 0$ and n odd



4. If $a_n < 0$ and n odd



The Leading Term Test implies that if you "zoom out" far enough, then all polynomials look like the power function made by the leading term of the polynomial. That is, if the exponent of the leading term is even, then the polynomial is U-shaped, and if the exponent of the leading term is odd, then the polynomial is shaped like $y = x^3$.

Zeros of a Polynomial Function

- Every polynomial function with degree $n \geq 1$ has at least one zero and at most n zeros (counting imaginary or complex zeros).

or

- Every polynomial function with degree $n \geq 1$ has *exactly* n zeros, if and only if the multiplicities are taken into account (again, counting imaginary zeros).

Example 4

Graph the polynomial function $f(x) = -3x^4 + 2x^3$.

Solution

Since the leading term here is $-3x^4$ then $a_n = -3 < 0$, and $n = 4$ even. Thus the end behavior of the graph as $x \to \infty$ and $x \to -\infty$ is that of Box #2, item 2.

We can find the zeros of the function by simply setting $f(x) = 0$ and then solving for x.

$$-3x^4 + 2x^3 = 0$$
$$-x^3(3x - 2) = 0$$

This gives

$$x = 0 \quad \text{or} \quad x = \frac{2}{3}$$

So we have two x–intercepts, at $x = 0$ and at $x = \frac{2}{3}$, with multiplicity $k = 3$ for $x = 0$ and multiplicity $k = 1$ for $x = \frac{2}{3}$.

To find the y–intercept, we find $f(0)$, which gives

$$f(0) = 0$$

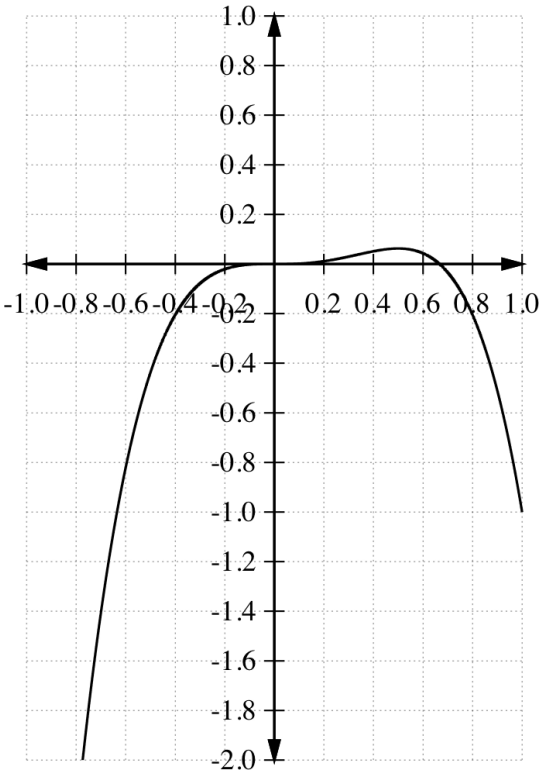
So the graph passes the y–axis at $y = 0$.

Since the x–intercepts are 0 and $\frac{2}{3}$, they divide the x–axis into three intervals: $(-\infty, 0), \left(0, \frac{2}{3}\right)$, and $\left(\frac{2}{3}, \infty\right)$. Now we are interested in determining at which intervals the function $f(x)$ is negative and at which intervals it is positive. To do so, we construct a table and choose a test value for x from each interval and find the corresponding $f(x)$ at that value.

Table 3.4:

Interval	**Test Value** x	$f(x)$	**Sign of** $f(x)$	**Location of points on the graph**
$(-\infty, 0)$	-1	-5	-	below the x–axis
$\left(0, \frac{2}{3}\right)$	$\frac{1}{2}$	$\frac{1}{16}$	+	above the x–axis
$\left(\frac{2}{3}, \infty\right)$	1	-1	-	below the x–axis

Those test points give us three additional points to plot: $(-1, -5), \left(\frac{1}{2}, \frac{1}{16}\right)$, and (1, -1). Now we are ready to plot our graph. We have a total of three intercept points, in addition to the three test points. We also know how the graph is behaving as $x \to -\infty$ and $x \to +\infty$. This information is usually enough to make a rough sketch of the graph. If we need additional points, we can simply select more points to complete the graph (Figure 10).



To summarize, the following procedure can be followed when graphing a polynomial function.

Graphing a Polynomial Function

1. Use the leading-term test to determine the end behavior of the graph.
2. Find the $x-$intercept(s) of $f(x)$ by setting $f(x) = 0$ and then solving for x.
3. Find the $y-$intercept of $f(x)$ by setting $y = f(0)$ and finding y.
4. Use the $x-$intercept(s) to divide the $x-$axis into intervals and then choose test points to determine the sign of $f(x)$ on each interval.
5. Plot the test points.
6. If necessary, find additional points to determine the general shape of the graph.

Example 5

Graph the polynomial function

$$g(x) = 2x^3 + 3x^2 - 50x - 75$$

Solution

Notice that the leading term is $2x^3$, where $n = 3$ odd and $a_n = 2 > 0$. This tells us that the end behavior will take the shape of item 3 in Box 2. Next we find the $x-$ and $y-$intercepts. Setting

$$g(x) = 0$$
$$2x^3 + 3x^2 - 50x - 75 = 0$$

We can factor this polynomial by grouping,

$$(2x + 3)(x^2) - (2x + 3)(25) = 0$$
$$(2x + 3)(x^2 - 25) = 0$$
$$(x + 5)(x - 5)(2x + 3) = 0$$

The zeros are -5, 5, and $\frac{-3}{2}$. And they divide the $x-$axis into four intervals:

$$(-\infty, -5) \qquad \left(-5, \frac{-3}{2}\right) \qquad \left(\frac{-3}{2}, 5\right) \qquad (5, \infty)$$

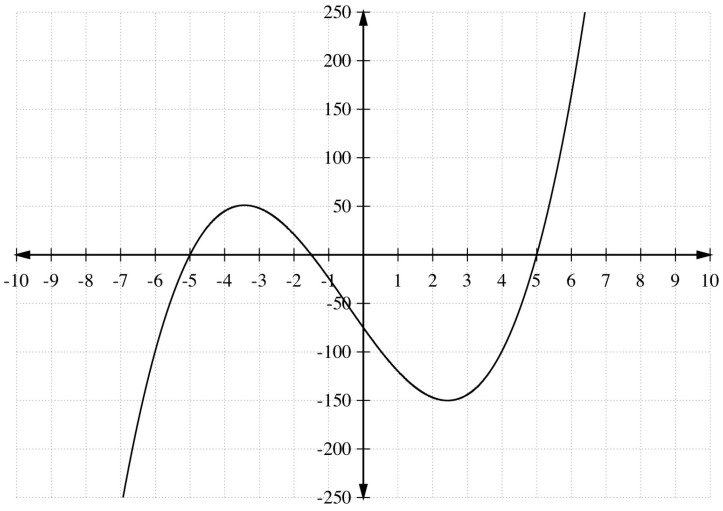
The $y-$intercept is when $y = g(0)$. Thus the graph intercepts the $y-$axis at $y = -75$.

Now we choose test points from each interval and find $g(x)$.

Table 3.5:

Interval	**Test value x**	**Value of $g(x)$ and its sign**	**Location of points on graph**
$(-\infty, -5)$	-6	-99	below the $x-$axis
$\left(-5, \frac{-3}{2}\right)$	-2	21	above the $x-$axis
$\left(\frac{-3}{2}, 5\right)$	0	-75	below the $x-$axis
$(5, \infty)$	6	165	above the $x-$axis

From the information obtained, we can roughly sketch the graph (below).



Applications, Technological Tools

You can use the same tools on a graphing calculator to analyze polynomials that you used to analyze quadratics. In particular, you can use the $Y =$ menu to graph a polynomial, use **WINDOW** and **ZOOM** to set up the view, and then use the **CALC** menu to find **MINIMUM, MAXIMUM,** and **ZEROs** of the function.

Two additional tools that are helpful for analyzing graphs with TI-83/84 calculators are the **TRACE** and **TABLE** functions.

If you have graphed a function using $Y =$ and set the **WINDOW** settings properly to allow you to see the graph, then you can use **TRACE** to "read" pairs of (x, y) coordinates on the graph.

((Insert screen shot of graph with the **TRACE** function turned on))

While this is useful, a drawback of the **TRACE** function is that when you use the arrow keys ($<$ and $>$) to "trace" along the function values you cannot easily control the amount by which the $x-$coordinate

changes each time you press the arrow key. One solution to find a particular value of the function is to press **TRACE** and then enter a number on the keypad. For example the sequence **TRACE** 5 will show the coordinates of the selected function when $x = 5$. (Note: this assumes that x is between **XMIN** and **XMAX**. If this is not the case, you will get an error).

((Insert screen shot of function with **TRACE** 5. It should show $X = 5$ in the lower left corner, and the corresponding $y-$value in the lower right))

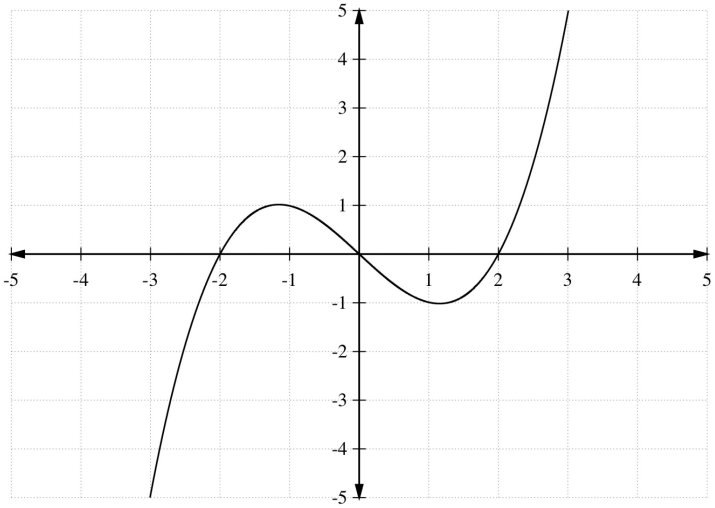
A third way to analyze a function is to use the table. If you enter **2nd GRAPH** then you see a table of function values.

((insert screen shot of table))

If you go to **TABLSET (2ND WINDOW)**, you can set the interval on the table, or change the independent variable to **ASK** and then type in $x-$values directly in the table.

Exercises

1. Sketch a graph of each power function using the properties of the power functions.
   (a) $f(x) = -3x^4$
   (b) $h(x) = \frac{1}{2}x^5$
   (c) $q(x) = 4x^8$
2. Consider the polynomial $f(x) = x^3 - 7x^2 + 10x$.
   (a) Find the zeros ($x-$intercepts) of $f(x)$.
   (b) Make a table to show on what intervals $f(x)$ is positive, and on what intervals $f(x)$ is negative
   (c) Sketch a graph of $f(x)$
3. Without graphing the function, what is the maximum number of roots and turning points of $g(x) = \frac{2}{3}x^5 - 4x^3 + 7x - \pi$. Justify your answer
4. Without graphing the function, what can you say about the end behavior of $g(x) = \frac{2}{3}x^5 - 4x^3 + 7x - \pi$? What happens as $x \to \infty$? What happens as $x \to -\infty$?
5. Use a graphing calculator to approximate the zeros of $k(x) = -3x^4 + 18x^2 - 5x + 3$
6. The graph of $r(x)$ is shown below



Sketch each of the following transformations of $r(x)$.

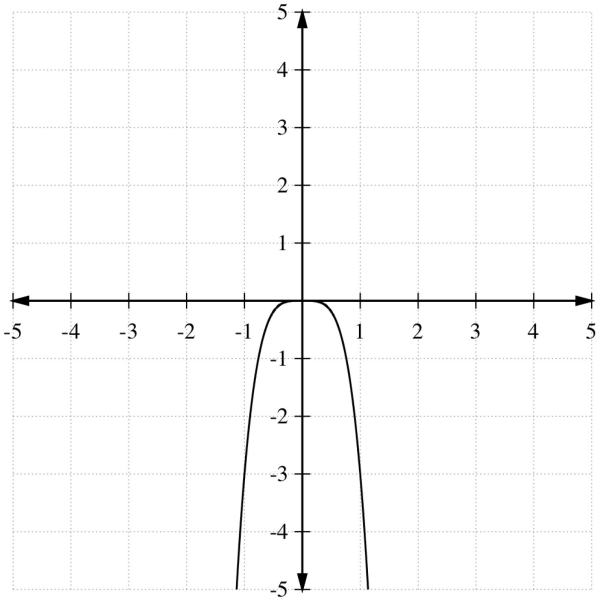
(a) $-r(x)$
(b) $r(x-3)$

(c) $r(-x) - 1$

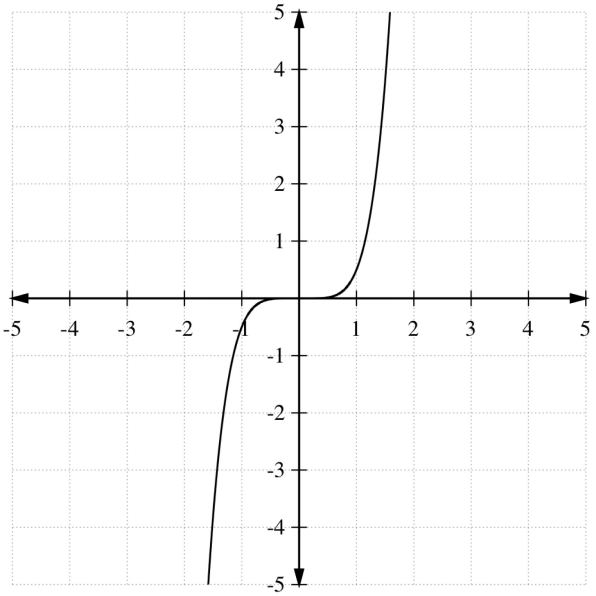
7. Which of the transformations of $r(x)$ in problem 6 have the same set of zeros as $r(x)$?

Answers

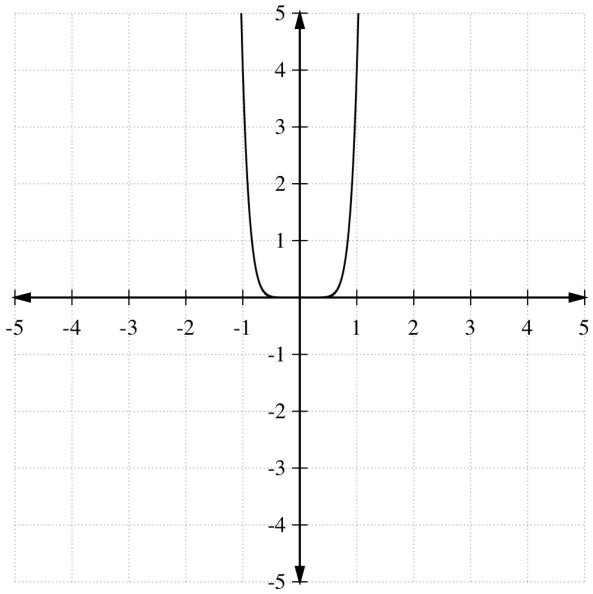
1. (a)



(b)



(c)

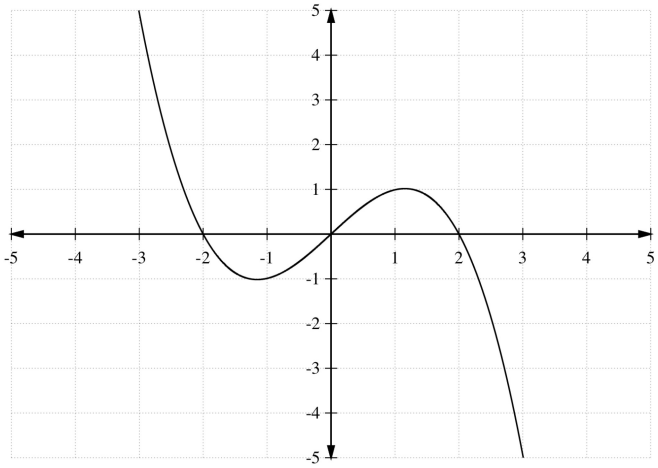


2. (a) $\{0, 5, 2\}$
   (b)

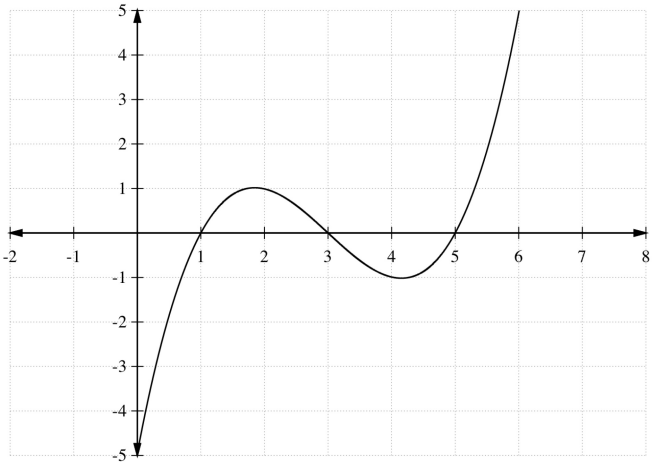
Table 3.6:

Interval	(	$0 < x < 2$	$2 < x < 5$	$x > 5$
Sign of $f(x)$	-	+	-	+

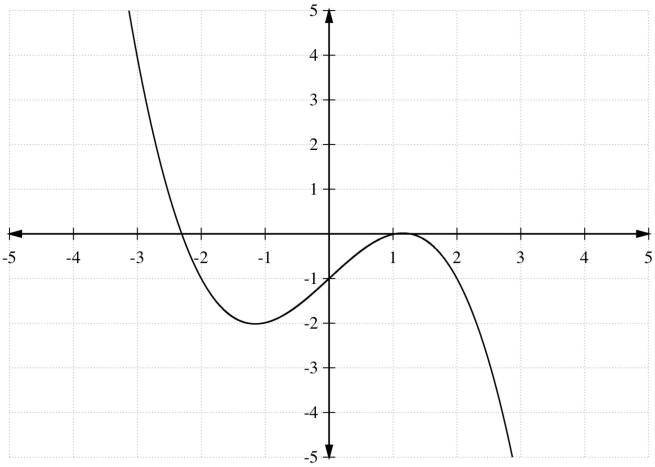
3. The maximum number of roots of $g(x)$ is 5 because the degree of $g(x)$ is 5. The maximum number of turning points is 4.
4. The end behavior of $g(x)$ is the same as the end behavior of x^5. This is because the leading term of $g(x)$ is $\frac{2}{3}x^5$.
5. $\{-2.6052, 2.33885\}$
6. (a)



(b)



(c)



7. Only the transformation in part (a), $-r(x)$ leaves the zeros the same. The other transformations involve vertical or horizontal shifts which change the $x-$ and $y-$intercepts.

3.3 Rational Functions

Learning objectives

- Find the $x-$ and $y-$intercepts of a rational function
- Find the horizontal, vertical, and oblique asymptotes
- Graph a rational function using $x-$ and $y-$intercepts and asymptotes

Standard Form of Rational Functions and the Domain of Rational Functions

Any function that has the form

$$f(x) = \frac{P(x)}{Q(x)}$$

where $P(x)$ and $Q(x)$ are polynomials and $Q(x) \neq 0$, is called a *rational function*. The *domain* of any rational function includes all real numbers x that do not make the denominator zero.

Example 1

What is the domain of $f(x) = \frac{1}{x}$?

Solution

Notice that the only input that can make the denominator equal to zero is $x = 0$. Thus we say that the domain of $f(x)$ is all real numbers except $x = 0$. When looking at the graph of $f(x) = \frac{1}{x}$ (Figure 13), you will notice that as x approaches 0 from the left, $f(x)$ decreases. But when x approaches 0 from the right, $f(x)$ increases. Because of this behavior, the $x-$axis and $y-$axis play the role of *horizontal* and *vertical asymptotes*, respectively.

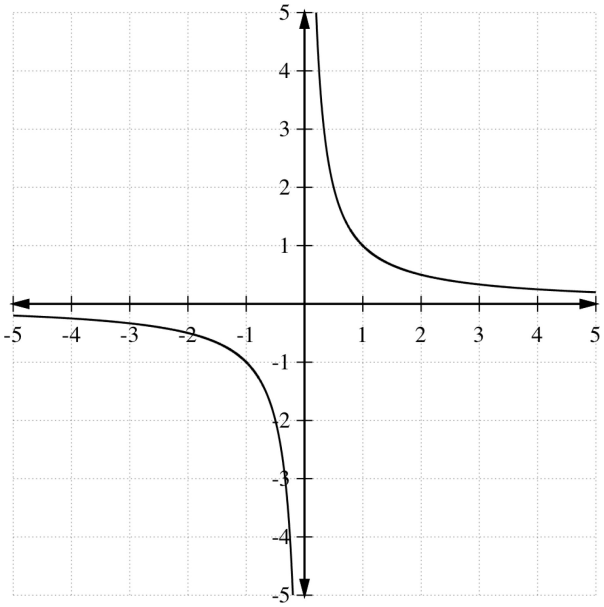
Graph Simple Rational Functions

Example 2

Graph the function $f(x) = \frac{1}{x}$.

Solution: We know that the domain of $f(x)$ is all real numbers excluding $x = 0$. The vertical line $x = 0$ is called a vertical asymptote. For $x < 0, f(x) < 0$, and for $x > 0, f(x) > 0$. Plotting a few sample points should indicate the shape of $f(x)$.

x	1	2	10	$\frac{1}{5}$	$\frac{1}{10}$	-1	$-\frac{1}{2}$	$-\frac{1}{10}$	-2	-10
$f(x) = \frac{1}{x}$	1	$\frac{1}{2}$	$\frac{1}{10}$	5	10	-1	-2	-10	$-\frac{1}{2}$	$-\frac{1}{10}$

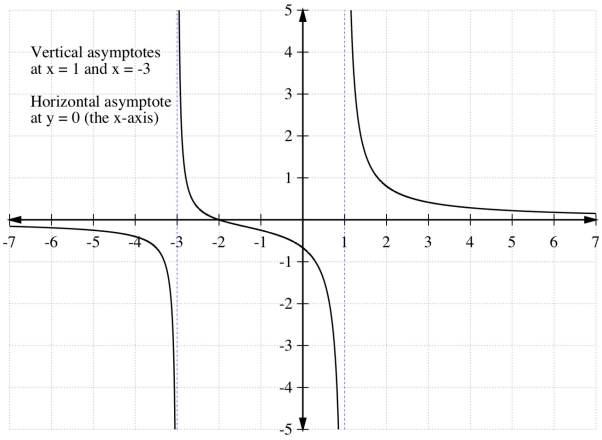


Example 3

What is the domain of $f(x) = \frac{x+2}{(x-1)(x+3)}$?

Solution

The domain is all real numbers except at the points that cause the denominator to equal zero, namely, at $x = 1$ and at $x = -3$. We will discuss this function in more detail in the next section. Below is a sketch of the graph of $f(x)$.



Vertical and Horizontal Asymptotes

An *asymptote* is a line or curve to which a function's graph draws closer without touching it. Functions cannot cross a vertical asymptote, and they usually approach horizontal asymptotes in their end behavior (i.e. as $x \to \pm\infty$). In the example above, the graph of $f(x) = \frac{1}{x}$ gets closer and closer to both axes, as x becomes large ($x \to \pm\infty$) or small ($x \to 0$), but it does not touch the axes. This behavior is typical for most rational functions.

There are three types of asymptotes: *horizontal, vertical* and *oblique*. We will analyze each one below.

Looking at the graph of $f(x) = \frac{x+2}{(x-1)(x+3)}$ from the previous example, you will notice that it has two *vertical asymptotes* (the vertical dotted lines), one is at $x = 1$ and the other is at $x = -3$.

We can find vertical asymptotes by simply equating the denominator to zero and then solving for x. In other words, if

$$f(x) = \frac{P(x)}{Q(x)}$$

Then setting $Q(x) = 0$, will give the vertical asymptote(s).

So if

$$f(x) = \frac{x+2}{(x-1)(x+3)}$$

setting

$$(x-1)(x+3) = 0$$

gives the vertical asymptotes at $x = 1$ and $x = -3$.

The horizontal asymptote a line parallel to the $x-$axis which the function approaches but does not reach as $x \to \infty$ and $x \to -\infty$. To find the horizontal asymptote, we follow the procedure below:

How to Find the Horizontal Asymptote

- Put the rational function in a standard form. That is, expand the numerator and denominator if they are written in a factored form.
- Remove all terms except the terms that contain the largest exponents of x in the numerator and the denominator.

- There are three possibilities:
  - If the degree of the numerator is smaller than the degree of the denominator, then the horizontal asymptote crosses the $y-$axis at $y = 0$. That is, it is the $x-$axis itself.
  - If the degree of the denominator and the numerator are the *same*, then the horizontal asymptote equals to the ratio of the leading coefficients.
  - If the degree of the numerator is larger than the degree of the denominator, then there is *no horizontal* asymptote.

Example 4

Find the vertical and horizontal asymptotes of

$$f(x) = \frac{2x^3 - 2x^2 + 5}{3x^3 - 81}$$

To find the vertical asymptote(s), set the denominator to zero and then solve for x.

$$\begin{aligned} 3x^3 - 81 &= 0 \\ 3x^3 &= 81 \\ x^3 &= 27 \\ x &= \sqrt[3]{27} \\ x &= 3 \end{aligned}$$

Thus the graph has a vertical asymptote at $x = 3$.

To find the horizontal asymptote, we follow the procedure above. Both the numerator and denominator are already written in standard form. Next, remove all terms except the largest exponents of x that are found in the numerator and the denominator):

$$\frac{2x^3}{3x^3}$$

Notice that the degree of the numerator and the denominator are the same and therefore the horizontal asymptote is the ratio of the coefficients,

$$\frac{2x^3}{3x^3} = \frac{2}{3}$$

So the horizontal asymptote is at $y = \frac{2}{3}$.

Example 5

Find the asymptotes of

$$f(x) = \frac{3x - 2}{2x^4 - 9}$$

Remove all terms except the leading terms,

$$\frac{3x}{2x^4}$$

Notice that the degree of the numerator is less than the degree of the denominator. Therefore, the horizontal asymptote is at $y = 0$, i.e., the $x-$axis plays the role of the horizontal asymptote. To find the vertical asymptote, set the denominator equal to zero and solve:

$$2x^4 - 9 = 0$$
$$x^4 = \frac{9}{4}$$
$$x = \pm\frac{\sqrt{6}}{2}$$

Example 6

Another example, consider the rational function

$$g(x) = \frac{2x^4 - 9}{3x - 2}$$

Removing all terms except the leading terms of the numerator and denominator,

$$\frac{2x^4}{3x}$$

Here, the degree of the numerator is larger than the degree of the denominator. Thus there is no horizontal asymptote.

Once you have found the asymptotes, it is relatively easy to graph rational functions. We will illustrate how to graph rational functions with two examples.

Example 7

Graph

$$T(x) = \frac{2x + 1}{x - 1}$$

Solution

Note that the domain of T is the set of all real numbers except 1, that is $x \neq 1$. This tells us that the line $x = 1$ is a vertical asymptote of $T(x)$. To graph T, there are four important items that you need to find: The $y-$intercept, the $x-$intercept, the vertical asymptote and the horizontal asymptote.

The $y-$intercept can be found by finding $y = T(0)$.

$$y = T(0) = \frac{2(0) + 1}{0 - 1} = -1$$

Thus the $y-$intercept is at point (0, -1).

The $x-$intercept can be found by setting $y = T(x) = 0$.

$$\frac{2x + 1}{x - 1} = 0$$

Note that a fraction $\frac{a}{b} = 0$ if and only if $a = 0$, so we can set the numerator of $T9$ solve

$$2x + 1 = 0$$
$$x = \frac{-1}{2}$$

Notice that we could have just set the numerator to zero and found the $x-$intercept. In general, set $P(x) = 0$ to find the $x-$intercept for any rational function. BUT, you must make sure that the $x-$value you find is still defined for the function. If both $P(x) = 0$ and $Q(x) = 0$ for some value of x then the graph has a hole in it.

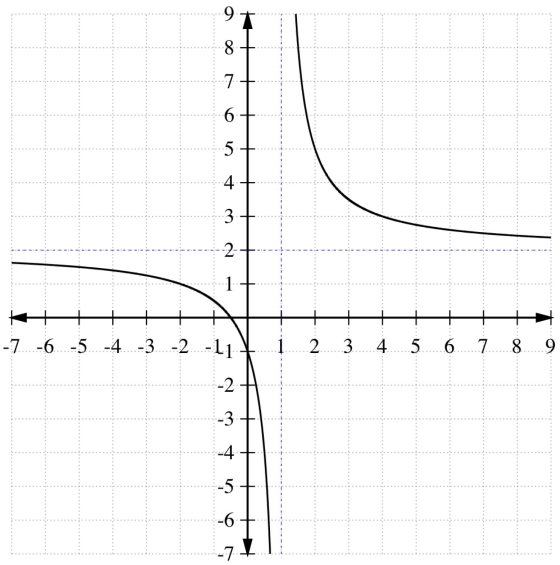
Next, the vertical asymptote. Set $Q(x) = 0$:

$$\begin{aligned} x - 1 &= 0 \\ x &= 1 \end{aligned}$$

And the horizontal asymptote:

$$\frac{2x+1}{x-1} \rightarrow \frac{2x}{x} = 2$$

Therefore, the vertical asymptote is at $x = 1$ and the horizontal asymptote is at $y = 2$. From this information, we can make the graph shown in Figure 15.



Example 8

Graph

$$g(x) = \frac{2x^2 + 1}{2x^2 - 3x}$$

Solution

The domain of g is the set of all real numbers except 0 and $\frac{3}{2}$, that is $\left\{x | x \neq 0 \text{ and } x \neq \frac{3}{2}\right\}$. The $y-$intercept is

$$y = g(0) = \frac{1}{0} = \text{undefined}$$

this tells us that there is no $y-$intercept. The $x-$intercept can be found by setting the numerator to zero,

$$2x^2 + 1 = 0$$
$$2x^2 = -1$$
$$x^2 = \frac{-1}{2}$$
$$x = \sqrt{\frac{-1}{2}}$$

Notice that this equation has no real solution and therefore, there is no x–intercept either.

The vertical asymptote can be found by setting the denominator to zero,

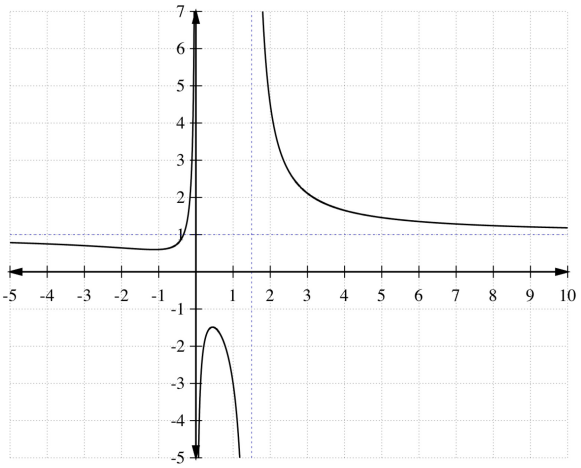
$$2x^2 - 3x = 0$$
$$x(2x - 3) = 0$$

The two solutions are $x = 0$ and $x = \frac{3}{2}$, and these are the vertical asymptotes.

Finally, the horizontal asymptote is found by analyzing the leading terms:

$$\frac{2x^2 + 1}{2x^2 - 3x} \rightarrow \frac{2x^2}{2x^2} = 1$$

That is, $y = 1$ is a horizontal asymptote. Again after substituting in some points, we can sketch the graph of $g(x)$ below.



Oblique Asymptotes

Thus far, we have restricted our discussion of rational functions to those where the degree of the numerator is less than or equal to the degree of the denominator. As our final analysis of graphing rational functions, we will consider the case when the degree of the numerator is greater than the degree of the denominator by one.

Example 9

Graph

$$g(x) = \frac{x^2 - 1}{x - 2}$$

Solution

First observe that the vertical asymptote is at $x = 2$. Notice that the degree of the numerator is greater than the degree of the denominator. We can change the form of the rational expression by long division. You may recall from algebra that polynomials can be divided (just like real numbers), and any rational functions can be written as

$$\frac{f(x)}{D(x)} = Q(x) + \frac{R(x)}{D(x)}$$

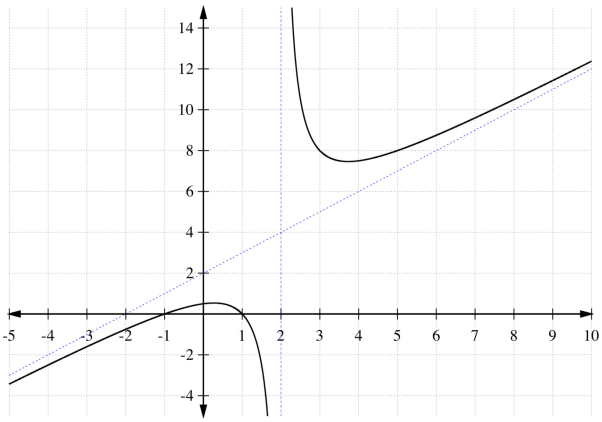
Doing the long division here,

$$\begin{array}{r} x+2 \\ x-2 \;\overline{)\,x^2+0x-1} \\ \underline{x^2-2x \quad \downarrow} \\ 2x-1 \\ \underline{2x-4} \\ 3 \end{array}$$

So in this case, the function $g(x)$ can be rewritten as

$$g(x) = \frac{x^2-1}{x-2} = x+2+\frac{3}{x-2}$$

The above equation tells us that as $x \to \pm\infty$, the graph of $g(x) = x+2+\frac{3}{x-2}$ gets closer and closer to the line $y = x+2$. Why? Suppose we let x be a big number, i.e. $x = 1,000,000$. Then the remainder of this rational function becomes $\frac{3}{999,998} \approx 0$ and we are left with $x+2$. We call this line an *oblique asymptote* and it is indicated by the dashed line in Figure 17.



Example 10

Graph

$$f(x) = \frac{x^2-x-2}{x-1}$$

Solution

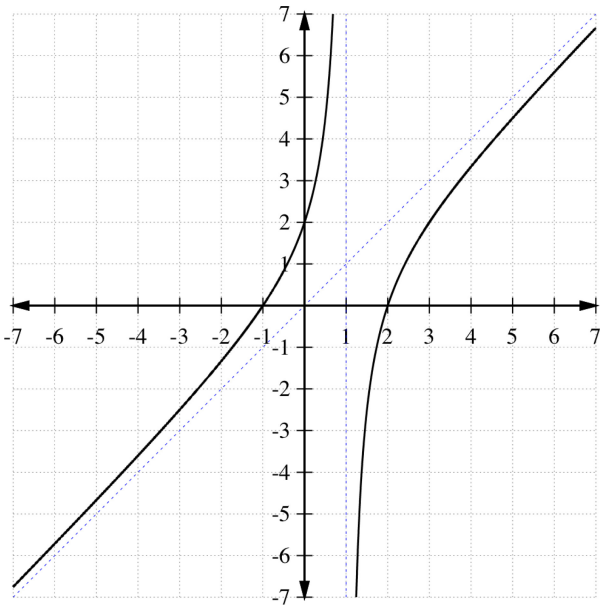
The vertical asymptote here is $x = 1$. Factoring the numerator and get

$$f(x) = \frac{x^2 - x - 2}{x - 1} = \frac{(x-2)(x+1)}{x-1}$$

Notice that the $x-$intercepts are at $x = 2$ and $x = -1$. By polynomial division, we get

$$f(x) = x - \frac{2}{x-1}$$

which indicates an oblique asymptote at $y = x$. Plotting few additional points, we finally obtain the graph shown below.



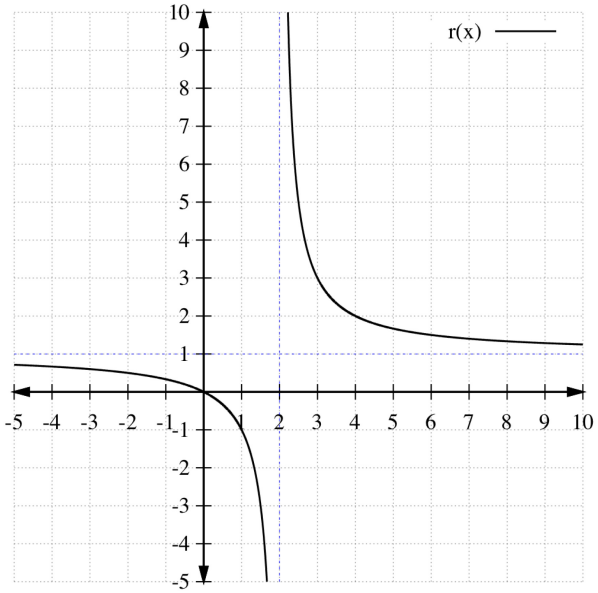
Graph Rational Functions Using Transformation

Just like polynomials, rational functions can be graphed using transformations. The main point to remember for graphing rational functions by transformations is that some transformations change the asymptotes while others do not.

- $r(x) + c$ is a vertical shift which moves each horizontal asymptote up c units (or down if $c < 0$).
- $r(x - c)$ is a horizontal shift which moves each vertical asymptote right c units (or left if $c < 0$).
- $a \cdot r(x)$ is a vertical stretch which moves horizontal asymptotes by a multiple of a (so this moves the horizontal asymptote closer to the $x-$axis if $a < 0$.
- $r(a \cdot x)$ is a horizontal compression which moves the vertical asymptotes closer to $y-$axis by a factor of $\frac{1}{a}$.
- $r(-x)$ is a reflection about the $y-$axis. All vertical asymptotes are also reflected.
- $-r(x)$ is a reflection about the $x-$axis. All horizontal asymptotes are also reflected.

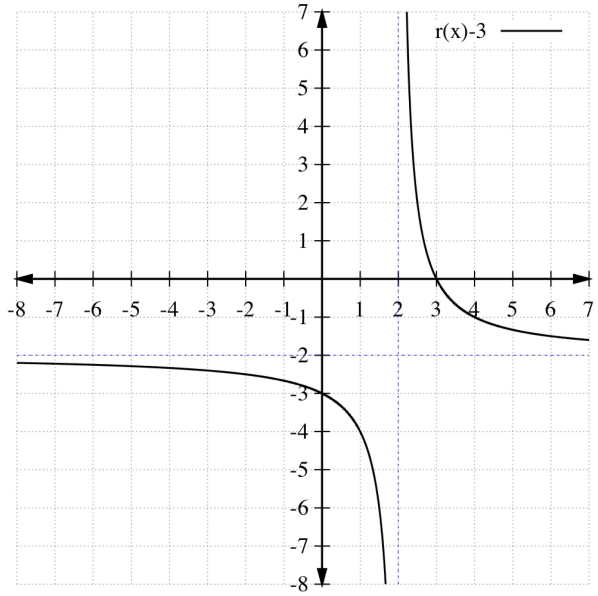
Example 11

A rational function $r(x)$ is shown in the figure below. Use the graph of $r(x)$ to sketch a graph of: a) $r(x) - 3$, b) $-r(x)$, c) $r(3 - x)$.

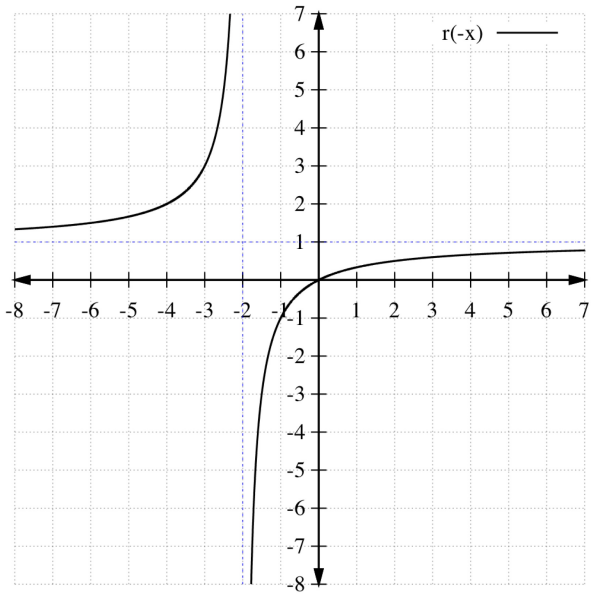


Solution

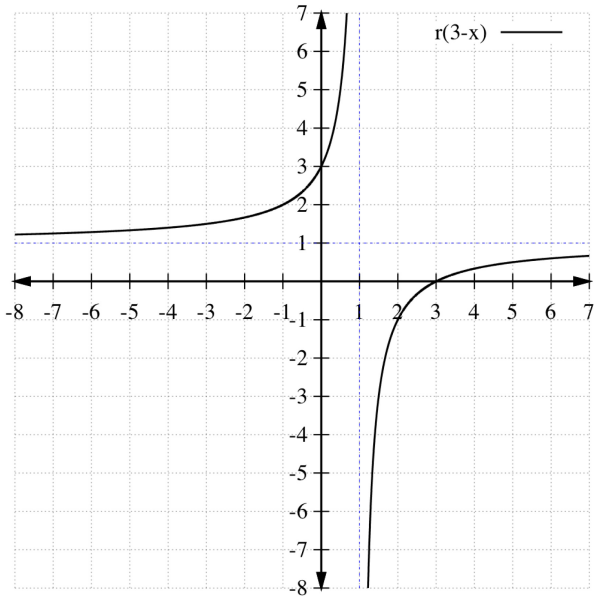
a) The horizontal asymptote moves down by three units



b) The function is reflected about the $x-$axis so the horizontal asymptote is also reflected:



c) $r(3-x) = r(-(x-3))$. First graph $r(-x)$, and then shift that graph three units to the right to get $r(-(x-3))$. The new vertical asymptote is $x = 1$.



Exercises

For each of the rational functions below, determine the domain, the asymptotes, the $x-$ and $y-$intercepts and then sketch the graph.

1. $f(x) = \frac{2x+5}{x-1}$
2. $f(x) = \frac{x+2}{x^2+1}$
3. $f(x) = \frac{9x^2-4}{3x-2}$
4. $f(x) = \frac{7}{3x^2}$
5. $f(x) = \frac{x^3}{x^3+1}$

6. $f(x) = \frac{14}{x^2-16}$
7. $f(x) = \frac{5(x-2)}{x^2-3x+2}$
8. $f(x) = \frac{x^2+1}{x-1}$
9. $f(x) = \frac{x^3-3x^2-4x}{x^2+3x}$
10. In physics, Boyle's law states that the product of the pressure P and the volume V of a container is always a constant. That is,

$$PV = \text{constant}$$

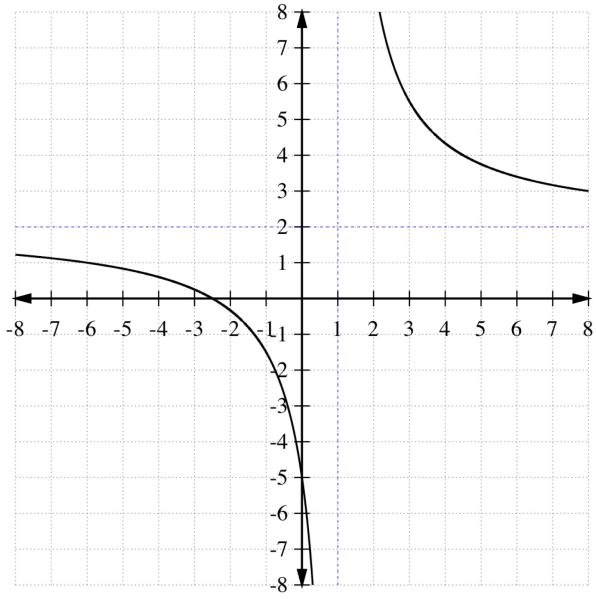
Suppose the constant is equal to 4000 $Pa.m^2$ (Pascals. square meters). So

$$P = \frac{4000}{V}$$

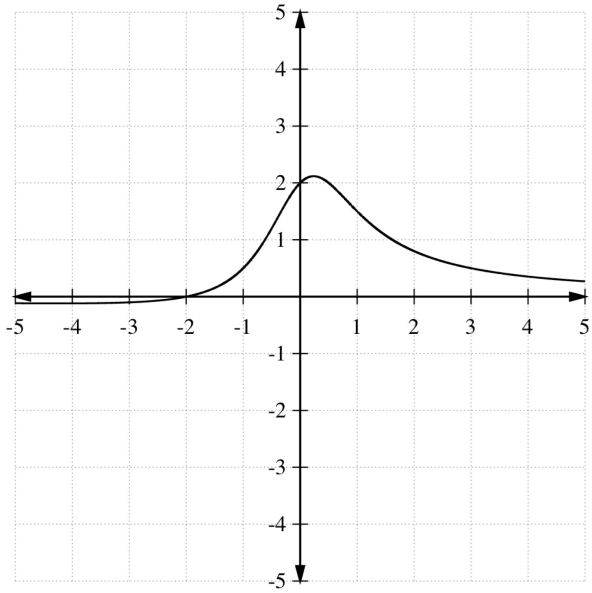
where the pressure is measured in Pascals and the volume is in meters-squared. Sketch the graph of the equation for $v > 0$.

Answers

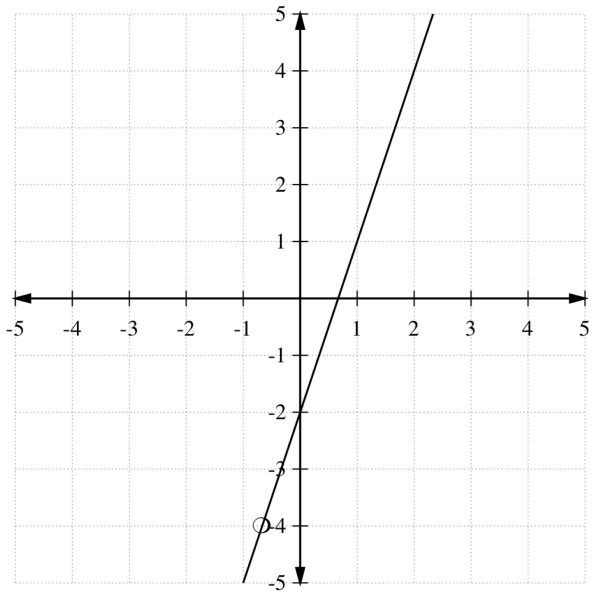
1. Domain: $x \neq 1$; Vertical asymptote $x = 1$; Horizontal asymptote $y = 2$



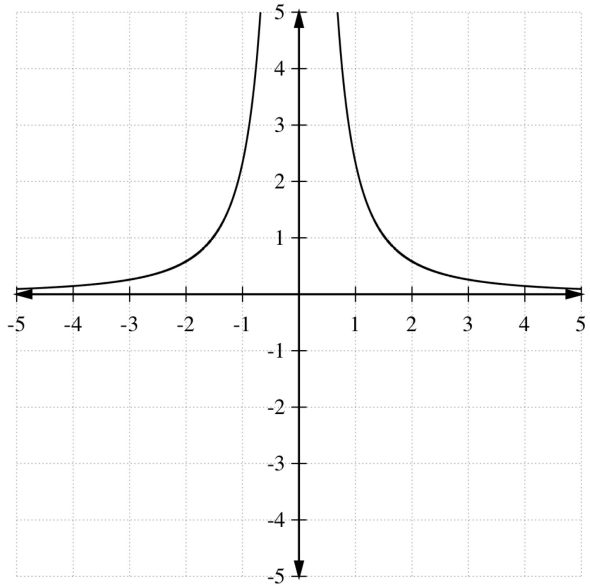
2. Domain: All real numbers; Vertical asymptote: none; Horizontal asymptote: $y = 0$



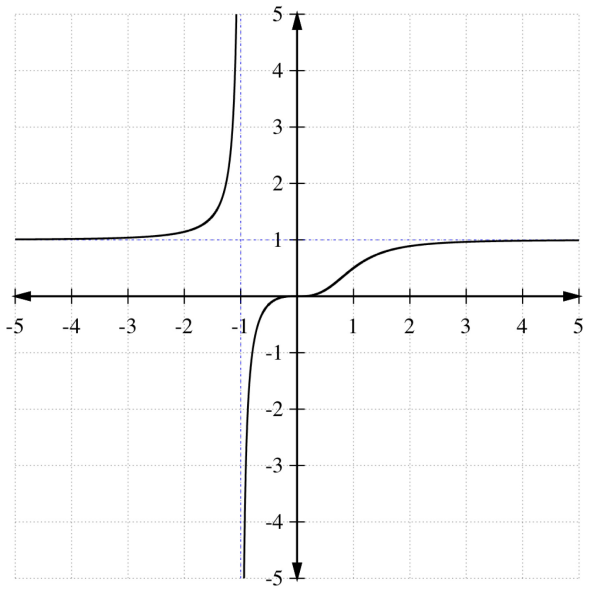
3. Domain: $x \neq \frac{2}{3}$; No asymptotes



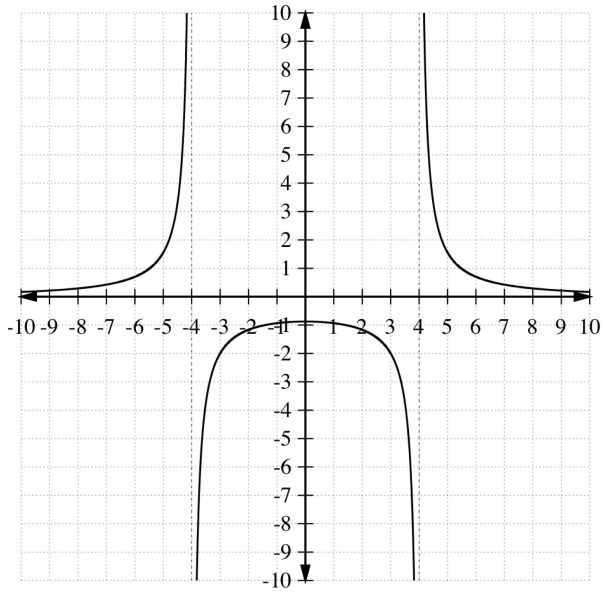
4. Domain: $x \neq 0$; Vertical asymptote $x = 0$; Horizontal asymptote $y = 0$



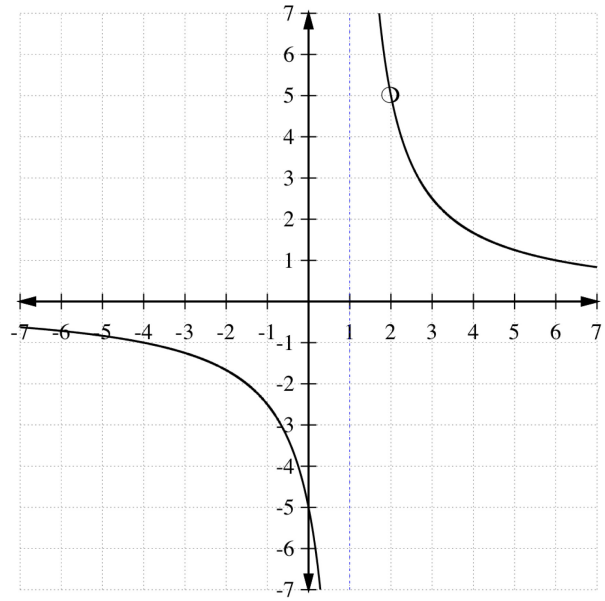
5. Domain: $x \neq -1$; Vertical asymptote $x = -1$; Horizontal asymptote $y = 1$



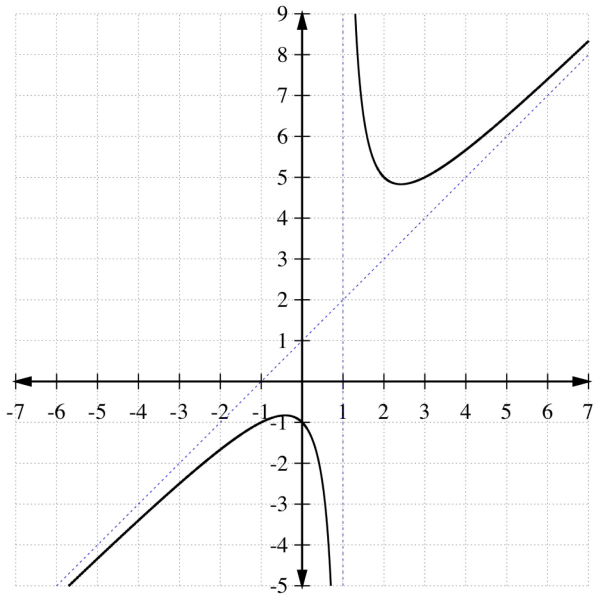
6. Domain: $x \neq \pm 4$; Vertical asymptote $x = \pm 4$; Horizontal asymptote $y = 0$



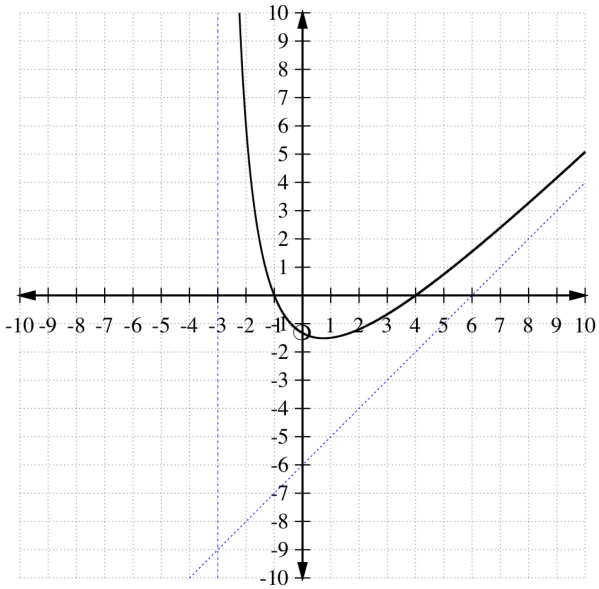
7. Domain: $x \neq 2, x \neq 1$; Vertical asymptote $x = 1$, Horizontal asymptote $y = 0$



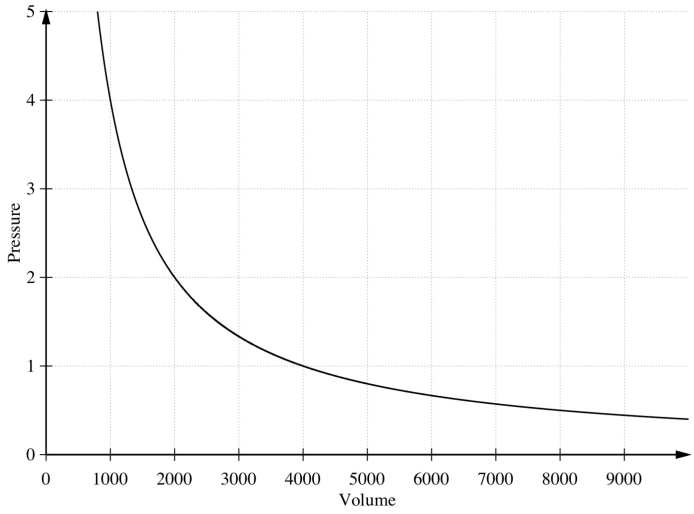
8. Domain: $x \neq 1$; Oblique asymptote $y = x + 1$; Vertical asymptote $x = 1$, Horizontal asymptote $y = 0$



9. Domain: $x \neq 0, x \neq -3$; Oblique asymptote $y = x - 6$; Vertical asymptote $x = -3$, No horizontal asymptote



10.



3.4 Analyzing Rational Functions

Learning Objectives

- Apply properties of rational functions to find roots and asymptotes
- Graph rational functions using roots, asymptotes, and behavior of the function around vertical asymptotes

Summarize Analysis of Rational Functions

Recall that *rational functions* are defined as $r(x) = \frac{p(x)}{q(x)}$ where $p(x)$ and $q(x)$ are polynomials. Rational functions can be slightly more complicated to analyze than polynomials, mostly due to the fact that anything divided by zero is undefined. In this section we review major points about analyzing rational functions, and show how a graphing calculator or other computer tool can be used to analyze rational functions.

Finding Vertical Asymptotes and Breaks in the Graph of a Rational Function

To find vertical asymptotes and breaks in the domain of a rational function, set the denominator equal to zero and solve for x. Given $r(x) = \frac{p(x)}{q(x)}$, set $q(x) = 0$ and solve for x.

Note, however, that some rational functions do not have vertical asymptotes and they are well-defined for all real numbers x. Other rational functions have a break in the function, but no vertical asymptote. This usually happens when one term in the numerator cancels with one term in the denominator. Below we illustrate examples of each of these cases.

$$r(x) = \frac{p(x)}{q(x)}$$

Example 1

Consider the following rational function. Find all restrictions on the domain and asymptotes

$$f(x) = \frac{x^2 + 2x - 35}{x - 5}$$

Solution

Factoring the numerator

$$f(x) = \frac{(x-5)(x+7)}{x+7}$$
$$f(x) = \frac{(x-5)\cancel{(x+7)}}{\cancel{x+7}}$$

Canceling

$$f(x) = x + 7, x \neq 5$$

Notice that there is no asymptote in this function, but rather a break in the graph at $x = 5$.

Example 2

Find the restrictions on the domain of

$$h(x) = \frac{3x}{x^2 - 25}$$

Solution

Setting the denominator equal to 0,

$$x^2 - 25 = 0$$
$$x^2 = 25$$
$$x = \pm 5$$

Thus, the domain of $h(x)$ is the set all all real numbers x with the restriction $x \neq \pm 5$. $h(x)$ has two vertical asymptotes, one at $x = 5$ and one at $x = -5$.

Example 3

Find the vertical asymptotes of

$$g(x) = \frac{x^3}{x^2 + 1}$$

Solution

Setting the denominator equal to zero,

$$x^2 + 1 = 0$$
$$x^2 = -1$$
$$x = \sqrt{-1}$$

There are no real solutions, so there are no vertical asymptotes and no restrictions on the domain of this function.

Example 4

What are the vertical asymptotes of the function

$$k(x) = \frac{x^2}{x^2 + 5}$$

Solution

Setting the denominator equal to zero,

$$x^2 + 5 = 0$$
$$x^2 = -5$$

Again there are no real solutions. The horizontal asymptote is $y = 1$.

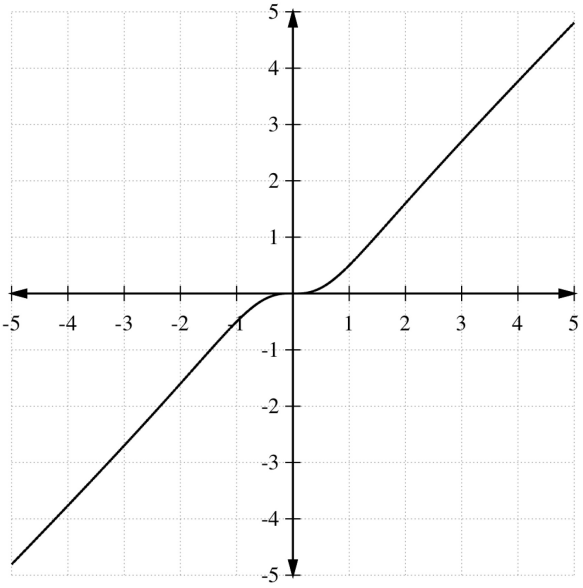
End Behavior of Rational Function

Previously you have analyzed the end behavior of polynomials. You may recall that, as a general rule, as you "zoom out," the graph of a polynomial looks like the graph of the power function made of the leading term of the polynomial. For instance, $g(x) = 3x^4 - 18x^3 - 15x$ looks like the function $y = 3x^4$ if you choose a sufficiently large range for x and y.

There is a similar rule for rational functions. The *End Behavior* of a rational function can often be identified by the horizontal asymptote. That is, as the values of x get very large or very small, the graph of the rational function will approach (but not reach) the horizontal asymptote. In Example 4, we could write as $x \to \infty, k(x) \to 1$ and as $x \to -\infty, k(x) \to 1$. This can be shortened by writing as $x \to \pm\infty, k(x) \to 1$.

Finding Oblique Asymptotes

Not all asymptotes of rational functions are vertical or horizontal. If we look at the graph of the rational function in Example 3 from above, $g(x) = \frac{x^3}{x^2+1}$, we can see that there is no horizontal asymptote of this function.



There is no horizontal asymptote in this function because the degree of the numerator is greater than the degree of the denominator.

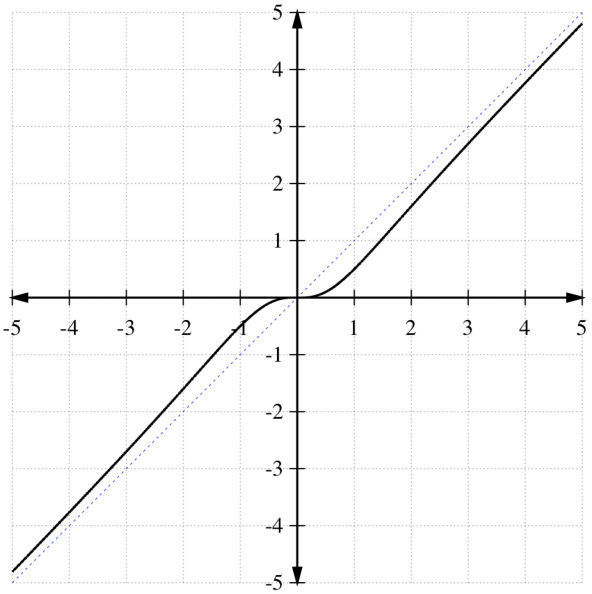
As a reminder, the following guidelines can help identify the asymptotes of a rational function $r(x) = \frac{f(x)}{D(x)}$:

- If the degree of the denominator is greater than the degree of the numerator, then the line $y = 0$ is a horizontal asymptote.
- If the degree of the numerator and the denominator are equal, then the line $y = \frac{a}{b}$ is a horizontal asymptote, where a is the leading coefficient of $f(x)$, the numerator, and b is the leading coefficient of $D(x)$, the denominator.
- If the degree of the numerator is larger than the degree of the denominator, then the quotient function, $Q(x)$, found by dividing the numerator and denominator of the rational function is an oblique asymptote. Recall that for any rational function $\frac{f(x)}{D(X)}$, you can use polynomial division to re-write that function in the form $\frac{f(x)}{D(x)} = Q(x) + \frac{R(x)}{D(x)}$ where $Q(x)$ is the quotient and $R(x)$ is the remainder.

To illustrate this last point, look at $g(x) = \frac{x^3}{x^2+1}$. By polynomial division we have,

$$\begin{array}{rll} & x & \leftarrow \text{Quotient} \\ x^2+1 \,\big) & \overline{x^3+0x^2+0x+0} & \leftarrow \text{Dividend} \\ & \underline{x^3 \qquad\quad +x} & \\ & \qquad\quad -x & \leftarrow \text{Remainder} \end{array}$$

So $g(x) = x - \frac{x}{x^2+1}$. This tells us that the line $y = x$ is an oblique asymptote of $g(x)$.



Notice that the oblique asymptotes of a rational function also describe the *end behavior* of the function. That is, as you "zoom out" from the graph of a rational function it looks like a line or the function defined by $Q(x)$ in $\frac{f(x)}{D(x)} = Q(x) + \frac{R(x)}{D(x)}$.

Example 5

Find the oblique asymptote of $f(x) = \frac{3x^3-5x^2+2}{x^2-3x}$. Sketch a graph of $f(x)$.

Solution

Using polynomial long division, $f(x) = 3x + 4 + \frac{4x^2}{x^2-3x}$. Thus, the line $y = 3x + 4$ is an oblique asymptote of $f(x)$.

To sketch the graph we can find the vertical asymptotes by setting the denominator equal to zero,

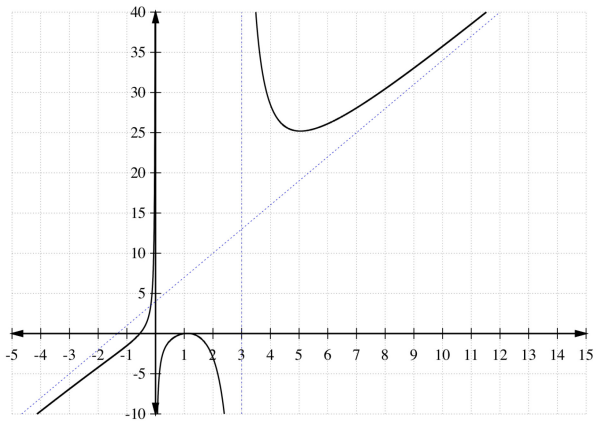
$$x^2 - 3x = 0$$
$$x(x-3) = 0$$

So the two vertical asymptotes are $x = 0$ and $x = 3$.

$f(0)$ is undefined, so there is no $y-$intercept. Also, there is no simple way to solve for the roots (setting the numerator equal to zero), but we can see by inspection that $f(1) = 0$. To get an idea of the shape of the graph we will make a table of a few test points. We used a calculator to evaluate decimal values of x in $f(x)$.

x	-2	-1	-0.1	1	2	4	3.5	5	6
$f(x)$	-4.2	-1.5	66.48	0	-3	28.5	39.6	25.2	26.1

Finally we use all of this information to make a sketch of the graph of $f(x)$:



Applications, Technological Tools

The last example illustrates the difficulty of graphing any rational function by hand. A graphing calculator can be a valuable tool to help with graphing and analyzing rational functions, provided you know how to use the calculator effectively. Below we illustrate a few of the benefits and potential pitfalls of graphing rational functions using TI-83/84 Calculators.

Most of the calculator analysis tools that we have reviewed for quadratic and polynomial functions also work with rational functions. You can use the $Y =$ and **GRAPH** menus to view the graph of a rational function. Once you have graphed a rational function you can use the **CALC** menu to find the roots and the min/max values of the function.

Potential Pitfalls of Graphing Rational Functions

There are two common pitfalls when using a calculator to graph a rational function:

1. When graphing a rational function by entering the function in the $Y =$ screen, remember that you need to use parenthesis to group the numerator and denominator of the rational function. Also,

the independent variable of a rational function is always X in the TI-83/84. Thus, for example, you would enter the function $f(t) = \frac{3t^2-2t+1}{7t^3+4}$ by typing "$\frac{(3X\text{^}2-2X+1)}{(7X\text{^}3+4)}$" in the $Y1 =$ slot.
2. Vertical asymptotes are sometimes graphed as vertical lines.
3. Graphs of rational functions can be difficult to interpret if the window settings are not chosen carefully.

The following example illustrates the second and third problems.

Example 6

Graph $f(x) = \frac{1}{x-3}$ on the window $[-10, 10] \times [-10, 10]$. (This means XMIN $= -10$, XMAX $= 10$, YMIN $= -10$, and YMAX $= 10$)

Solution

((insert screen shot of the graph of $y = \frac{1}{x-3}$ showing that vertical line at $x = 3$))

$f(x)$ is undefined and has a vertical asymptote at $x = 3$, but the way the graphing calculator draws the graph, it shows a vertical line at $x = 3$. One way to "fix" this problem is to press **MODE** and select the option "Dot" rather than "Connected". However, dot graphs can be hard to interpret as well.

Using the Table to Examine Behavior of a Rational Function

Despite the potential pitfalls of using a graphing calculator to graph rational functions, the **TABLE** function of the graphing calculator is one very effective tool for analyzing the behavior of a rational function near its vertical asymptotes. Let's consider the function from Example 5 above, $f(x) = \frac{3x^3-5x^2+2}{x^2-3x}$. The vertical asymptotes of $f(x)$ are $x = 3$ and $x = 0$. To analyze the behavior of $f(x)$ near these asymptotes, first enter the equation of $f(x)$ in the $Y =$ menu. Next, press **2ND WINDOW** to go to the **TABLSET** menu. There, change the setting of Independent to **ASK**.

((Calculator Image: Show **TABLE SETUP** menu))

Press **2ND GRAPH** to enter the **TABLE** screen.

In the table you can enter x values in the first column, and the second column will display corresponding y values.

To analyze the behavior of $f(x)$ near $x = 0$, enter several values for x getting progressively closer to 0.

x	-1	-0.5	-0.25	-0.1	-0.01	-0.001
$f(x)$	-1.5	0.214	2.019	6.281	66.428	666.443

It looks like as $x \to 0$ from the left that $f(x)$ becomes large. (In calculus you will use more rigorous language to describe this situation). Let's look at x values larger than zero, that is look at how $f(x)$ behaves as $x \to 0$ from the right.

x	1	0.5	0.25	0.1	0.01	0.001
$f(x)$	0	-0.9	-2.523	-6.734	-66.873	-666.887

This shows that as $x \to 0$ from the right, $f(x) \to -\infty$.

Likewise, you can use the table to show that as $x \to 3$ from the left, $f(x) \to -\infty$, and as $x \to 3$ from the right, $f(x) \to \infty$.

While using the table to sketch a graph by hand may be redundant (after all, you have a graphing calculator!), this information helps you interpret the output of the graphing tool, which is an important skill to develop as you learn to use a graphing calculator.

Exercises

For each function in questions 1-5, describe the behavior of the rational function.

a) What are its asymptotes?

b) What is the end behavior of the function?

c) How does the function behave as x approaches each vertical asymptote form the left and from the right?

d) Sketch a graph of the function an an appropriate domain

1. $g(x) = \frac{4x^3-7x}{3x^3-x}$
2. $h(x) = \frac{x-2}{x+1}$
3. $k(x) = \frac{5}{x^2-12}$
4. $f(x) = \frac{2x^3+2x^2-1}{x^2-5x+4}$
5. $p(x) = \frac{3x^2-3x}{2x^2+1}$

Answers

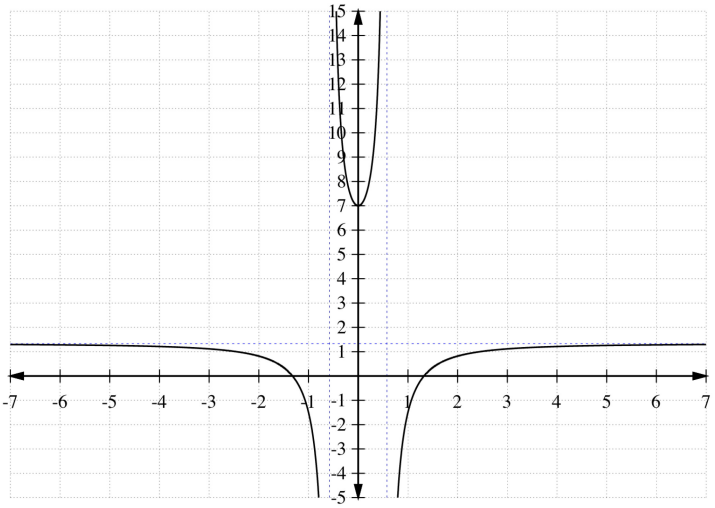
1. Notice that $g(x) = \frac{4x^3-7x}{3x^3-x} = \frac{x(4x^2-7)}{x(3x^2-1)} = \frac{4x^2-7}{3x^2-1}$, $x \neq 0$.

   (a) Vertical asymptotes at $x = \pm\sqrt{\frac{1}{3}}$. Horizontal asymptote $y = \frac{4}{3}$. $g(x)$ is undefined at $x = 0$, but there is not an asymptote there.

   (b) As $x \to \pm\infty, g(x) \to \frac{4}{3}$.

   (c) As $x \to -\sqrt{\frac{1}{3}}$ from the left $g(x) \to -\infty$ and as $x \to -\sqrt{\frac{1}{3}}$ from the right, $g(x) \to \infty$. As $x \to \sqrt{\frac{1}{3}}$ from the left, $g(x) \to \infty$ and as $x \to \sqrt{\frac{1}{3}}$ from the right, $g(x) \to -\infty$.

   (d)

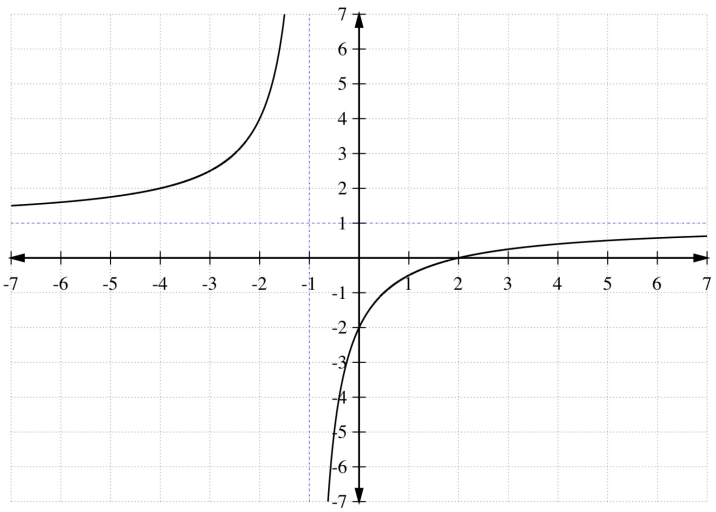


2. (a) Vertical asymptote at $x = -1$. Horizontal asymptote $y = 1$.

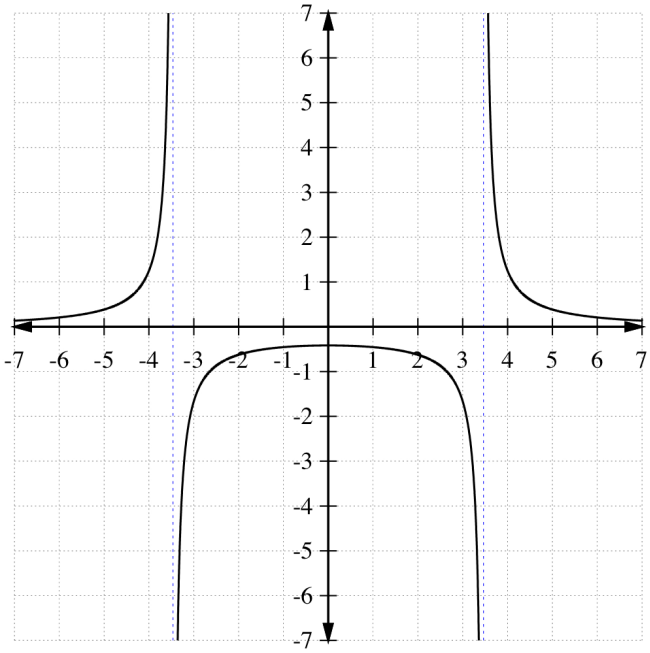
   (b) As $x \to \pm\infty, h(x) \to 1$.

   (c) As $x \to -1$ from the left $h(x) \to \infty$ and as $x \to -1$ from the right, $h(x) \to -\infty$.

(d)

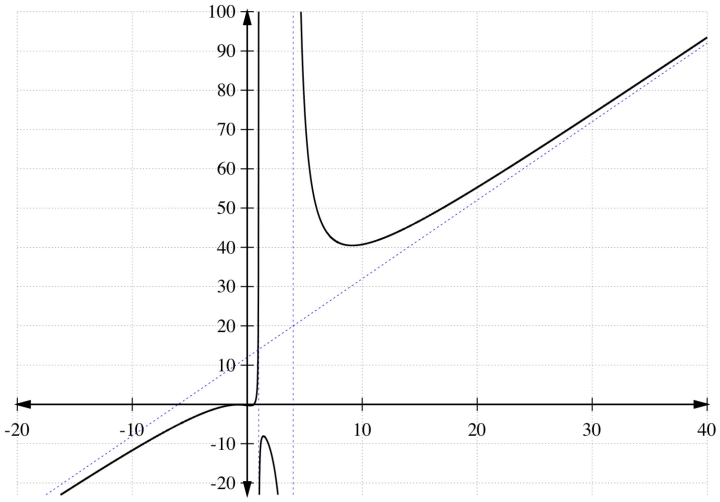


3. (a) Vertical asymptotes at $x = \pm 2\sqrt{3}$. Horizontal asymptote is $y = 0$.
   (b) As $x \to \pm\infty, k(x) \to 0$.
   (c) As $x \to -2\sqrt{3}$ from the left $k(x) \to \infty$ and as $x \to -2\sqrt{3}$ from the right, $k(x) \to -\infty$. As $x \to 2\sqrt{3}$ from the left $k(x) \to -\infty$ and as $x \to 2\sqrt{3}$ from the right, $k(x) \to \infty$.
   (d)



4. (a) Vertical asymptotes at $x = 4$ and $x = 1$. Oblique asymptote $y = 2x + 12$.
   (b) As $x \to -\infty, f(x) \to -\infty$. As $x \to \infty$, $f(x) \to \infty$.
   (c) As $x \to 1$ from the left $f(x) \to \infty$ and as $x \to 1$ from the right, $f(x) \to -\infty$. As $x \to 4$ from the left $f(x) \to -\infty$ and as $x \to 4$ from the right, $f(x) \to \infty$

(d)



5. (a) No vertical asymptotes. horizontal asymptote at $y = \frac{3}{2}$.
   (b) As $x \to \pm\infty, p(x) \to \frac{3}{2}$.
   (c) N/A: there are no vertical asymptotes.

3.5 Polynomial and Rational Inequalities

Learning Objectives

- Understand the properties of quadratic inequalities.
- Understand the properties of polynomial inequalities.
- Understand the properties rational inequalities.
- Apply understanding of quadratic, polynomial, and rational inequalities to draw graphs.

Quadratic Inequalities

Quadratic inequalities are inequalities that have one of the following forms

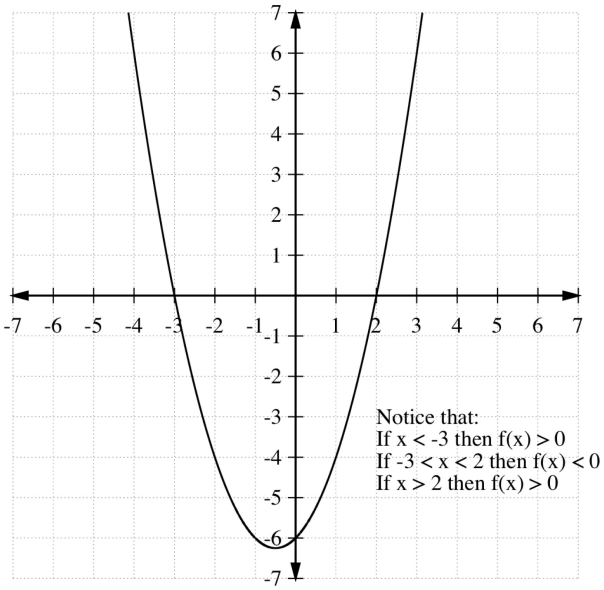
$$ax^2 + bx + c > 0$$

and

$$ax^2 + bx + c < 0$$

We can solve these inequalities by using the techniques that we have learned about solving quadratic equations. For example, consider the graph of the equation:

$$y = f(x) = x^2 + x - 6$$



Notice that the curve intersects the $x-$axis at -3 and 2. From graph, we notice the followings

- If $x < -3$ then $f(x) > 0$
- If $-3 < x < 2$, then $f(x) < 0$
- If $x > 2$, then $f(x) > 0$

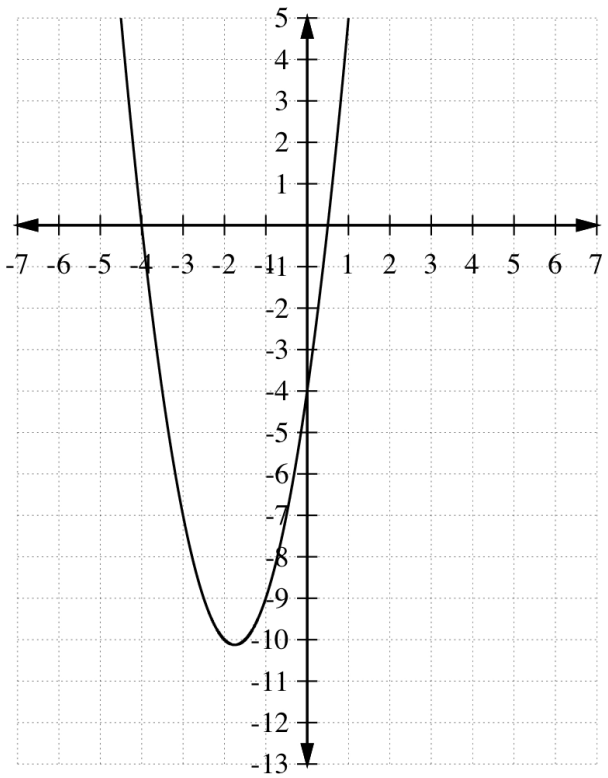
Therefore, $x^2 + x - 6 > 0$ whenever $x < -3$ or $x > 2$, and $x^2 + x - 6 < 0$ when $-3 < x < 2$.

Example 1

What is the solution set of the inequality $2x^2 + 7x - 4 < 0$?

Solution

It is best to graph the function $f(x) = 2x^2 + 7x - 4$ and look for the the values of x such that the inequality $f(x) < 0$ is true.



Thus from graph, $2x^2 + 7x - 4 < 0$ only if

$$-4 < x < \frac{1}{2}$$

So the solution set is $x \in \left(-4, \frac{1}{2}\right)$ or in set builder notation, $\left\{x | -4 < x < \frac{1}{2}\right\}$.

Although the method of graphing to find the solution set of an inequality is easy to follow, another algebraic method can be used. The algebraic method involves finding the $x-$intercepts of the graph and then dividing the $x-$axis into intervals separated by the $x-$intercepts. The examples below illustrate the method.

Example 2

Find the solution set of the quadratic inequality $x^2 + 2x - 8 > 0$.

Solution

As you know the graph of the function will immediately tell us at which intervals the inequality is positive. However, we will not draw a graph and find the solution without its aid. Rather, we solve

$$x^2 + 2x - 8 = 0$$
$$(x + 4)(x - 2) = 0$$

The two solutions to this quadratic equation are $x = -4$ and $x = 2$, thus, the $x-$intercepts of the function $f(x) = x^2 + 2x - 8$ are -4 and 2. These points divide the $x-$axis into three intervals: $(-\infty, -4), (-4, 2)$ and $(2, \infty)$. We can choose a test point from each interval and substitute it into $f(x)$ and see if it is negative or positive. This procedure can be simplified by making the table shown below.

Table 3.7:

Interval	**Test Point**	$x^2 + 2x - 8$ **Positive/Negative?**	**Part of Solution set?**
$(-\infty, -4)$	-5	+	yes
(-4, 2)	1	-	no
$(2, +\infty)$	3	+	yes

From the table, we conclude that since $x^2 + 2x - 8 > 0$ if and only if $x < -4$ and $x > 2$. The solution set can also be written as

$$x \in (-\infty, -4) \cup (2, +\infty)$$

Some problems in science involve quadratic inequalities. The example below illustrates one such application.

Example 3

A rectangle has a length 10 meters more than twice the width. Find the possible widths such that the area of the rectangle can not exceed 100 squared meters.

Solution

Let W be the width of the rectangle and L its length. Thus

$$\begin{aligned} L &= 10 + 2W \\ W &= W \end{aligned}$$

The area of a rectangle is

$$\begin{aligned} A &= LW \\ &= (10 + 2W)(W) \\ &= 10W + 2W^2 \end{aligned}$$

We are requiring that the area can not exceed 100 m^2. So

$$10W + 2W^2 < 100$$

or

$$2W^2 + 10W - 100 < 0$$

Dividing both sides by 2,

$$W^2 + 5W - 50 < 0$$

Finding the roots for $W^2 + 5W - 50 = 0$,

$$W^2 + 5W - 50 = (W + 10)(W - 5)$$

So the partition points are 5 and -10 and thus we have three intervals. Since width can not be negative, we can safely ignore -10. The maximum area is 100 m^2 and thus the width must be $W < 5$.

Solve Polynomial Inequalities

Solving polynomial inequalities is very similar to solving quadratic inequalities. The basic steps are the same:

1. Set up the inequality in the form $p(x) > 0$ (or $p(x) < 0, p(x) \leq 0, p(x) \geq 0$)
2. Find the solutions to the equation $p(x) = 0$.
3. Divide the number line into intervals based on the solutions to $p(x) = 0$.
4. Use test points to find solution sets to the equation.

Example 4

Solve $x^3 - 3x^2 + 2x \geq 0$

Solution

The polynomial is already in the correct form $p(x) \geq 0$ so we solve the equation

$$x^3 - 3x^2 + 2x = 0$$
$$x(x^2 - 3x + 2) = 0$$
$$x(x-2)(x-1) = 0$$

The zeros are at $x = 0, x = 1$, and $x = 2$.

Table 3.8:

Interval	Test Point	Positive/Negative?	Part of Solution set?
$(-\infty, 0)$	-5	-	no
(0, 1)	$\frac{1}{2}$	+	yes
(1, 2)	$\frac{3}{2}$	-	no
$(2, +\infty)$	3	+	yes

Notice that this inequality is greater than or equal to zero, so we include the zeros in the solution set. Therefore the solutions are $x \in [0, 1] \cup [2, +\infty]$.

Example 5

Solve $6x^4 + 20x^2 < 25$.

Solution

First we will change the inequality to $6x^4 + 20x^2 - 25 < 0$. Now, solve the equation $6x^4 + 20x^2 - 25 = 0$.

$$6x^4 + 20x^2 - 25 = 0$$
$$(3x^2 - 5)(2x^2 + 5) = 0$$

The first term yields the solutions $x = \pm\sqrt{\frac{5}{3}}$ and there are no real solutions for the second term.

Table 3.9:

Interval	Test Point	Positive/Negative?	Part of Solution set?
$\left(-\infty, -\sqrt{\frac{5}{3}}\right)$	-3	+	no
$\left(-\sqrt{\frac{5}{3}}, \sqrt{\frac{5}{3}}\right)$	0	-	yes
$\left(\sqrt{\frac{5}{3}}, +\infty\right)$	3	+	no

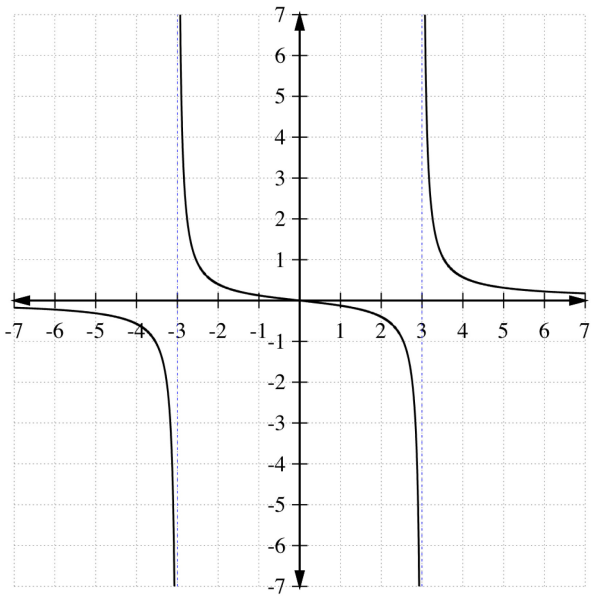
Finally, the solution set is $x \in \left(-\sqrt{\frac{5}{3}}, \sqrt{\frac{5}{3}}\right)$.

Notice that all of the polynomials we have picked so far are easily factorable. At this point the set of polynomial inequalities you can solve using algebraic methods is fairly limited. Later in this section we look at how a graphing calculator can help you solve polynomial inequalities. Also, once you learn more advanced methods for solving polynomial equations and finding rational zeros, you can apply those methods to solving different types of polynomial inequalities.

Solve Rational Inequalities

The method outlined above also works for solving inequalities involving rational functions. Recall that for rational functions you can find the roots (or zeros) by setting the numerator equal to zero.

However there is one step added to the process of solving rational inequalities because a rational function can also change signs at its vertical asymptotes or at a break in the graph. For instance, look at the graph of the function $r(x) = \frac{x}{x^2-9}$ below.



If we want to solve the inequality $\frac{x}{x^2-9} > 0$, then we need to use the following critical points: $x = 0, x = 3$, and $x = -3$. $x = 0$ is the solution of setting the numerator equal to 0, and this gives us the only root of the function. $x = \pm 3$ are the vertical asymptotes, the $x-$coordinates that make the function undefined because putting in 3 for x will cause a division by zero.

Using the graph or test points, we can build the table,

Table 3.10:

Interval	Test Point	Positive/Negative?	Part of Solution set?
$(-\infty, -3)$	-4	-	no
(-3, 0)	-2	+	yes
(0, 3)	2	-	no
$(3, +\infty)$	4	+	yes

Thus, the solutions to $\frac{x}{x^2-9} > 0$ are $x \in (-3, 0) \cup (3, +\infty)$.

Example 6

Find the solution set of the inequality

$$\frac{4x-12}{3x-2} < 0$$

From the numerator we solve $4x - 12 = 0$ or $x = 3$. In the denominator, solve $3x - 2 = 0$ and we find the critical point $x = \frac{2}{3}$.

Making the table

Table 3.11:

Interval	Test Point	Positive/Negative?	Part of Solution set?
$\left(-\infty, \frac{2}{3}\right)$	0	+	no
$\left(\frac{2}{3}, 3\right)$	1	-	yes
$(3, +\infty)$	5	+	no

Therefore, the solution set includes the numbers in the interval $\left(\frac{2}{3}, 3\right)$. Or in set-builder notation, the solution is $\left\{x | \frac{2}{3} < x < 3\right\}$.

Summarize Polynomial and Rational Inequalities

The basic method for solving polynomial and rational inequalities is the same:

- Re-write the inequality so that it is in terms of zero.
- Find the zeros and critical points to divide the domain into intervals
- Use test points in each interval to see which intervals satisfy the inequality
- Build your solution set from the table of intervals and test points

Example 7

Solve the inequality $\frac{3x+2}{x^2} > 3$

Solution

First, we use algebra to re-write the inequality to get zero on one side

$$\frac{3x+2}{x^2} > 3$$
$$\frac{3x+2}{x^2} - 3 > 0$$
$$-\frac{3x^2+3x+2}{x^2} > 0$$
$$\frac{3x^2-3x-2}{x^2} < 0$$

Notice in the last step we multiplied both sides by -1, so we changed the direction of the inequality. The numerator cannot be factored, so we use the quadratic formula to solve $3x^2 - 3x - 2 = 0$.

$$x = \frac{3 \pm \sqrt{9 - 4(3)(-2)}}{6}$$
$$x = \frac{3 \pm \sqrt{33}}{6}$$

So the two zeros of the rational function are $x = \frac{3+\sqrt{33}}{6} \approx 1.457$ and $x = \frac{3-\sqrt{33}}{6} \approx -0.457$. The final critical point is $x = 0$ because of the x^2 term in the denominator of the inequality. Using these values we construct the table:

Table 3.12:

Interval	**Test Point**	**Positive/Negative?**	**Part of Solution set?**
$\left(-\infty, \frac{3-\sqrt{33}}{6}\right)$	-1	+	no
$\left(\frac{3-\sqrt{33}}{6}, 0\right)$	$-\frac{1}{4}$	-	yes
$\left(0, \frac{3+\sqrt{33}}{6}\right)$	1	-	yes
$\left(\frac{3-\sqrt{33}}{6}, +\infty\right)$	2	+	no

The final solution set is $x \in \left(\frac{3-\sqrt{33}}{6}, 0\right) \cup \left(0, \frac{3+\sqrt{33}}{6}\right)$

Notice in this example that the intervals of the solution are on both sides of $x = 0$. You may be tempted to include $x = 0$ in the solution set, but that will not work. The reason is because at $x = 0$ the rational function is undefined, so 0 cannot satisfy the inequality.

Applications, Technological Tools

Solving quadratic, polynomial, and rational inequalities is much easier with a calculator. Two specific functions that a TI-83/84 calculator provides to help solve rational inequalities are:

1. Using the calculator to graph a function and using the **CALC** menu to identify its roots.
2. Using the table function to substitute test values into the function.

In a previous section we outlined step-by-step directions for doing both of these tasks. Here we will look at one example that brings these steps together to solve a polynomial inequality that could not be done using simple algebraic methods.

Example 8

The McNeil Surf Company makes wetsuits. For a given number of wetsuits x, McNeil's profit, in dollars, is given by the function $P(x) = -0.01x^2 + 25x - 3000$.

a) If the manager of McNeil wants the profit to stay above \$9,000, what is the minimum and maximum number of wetsuits they can manufacture to maintain that level of profit?

b) What is the maximum profit McNeil can make?

c) Can you explain why this shape might make sense for a profit function?

Solution

a) Set up the inequality

$$-0.01x^2 + 25x - 3000 > 9000$$
$$-0.01x^2 + 25x - 12000 > 0$$

With a calculator you can graph the function $Y_1 = -0.01x^2 + 25x - 12000$. Here is the graph using a window $[-1000, 4000] \times [-5000, 15000]$.

((**INSERT** TI-83 Graph with **WINDOW**, $Xmin = -1000, Xmax = 4000, Xscl = 500, Ymin = -5000, Ymax = 5000, Yscl = 1000$ $xres = 1$))

((**INSERT** TI-83/84 Graph of the function))

Using the **CALC** menu **(2ND TRACE)**, and selecting the option **ZEROS**, we can see that the zeros of $Y_1 = -0.01x^2 + 25x - 12000$ are at $x = 647.920$ and $x = 1852.080$. By inspecting the graph, we can see that the solution set to the inequality $-0.01x^2 + 25x - 12000 > 0$ is $x \in (648, 1852)$.

b) Keeping the same graph open, use **CALC MAXIMUM** to solve for the maximum profit. The maximum is at (1250, 3625), indicating that the maximum profit is \$3625 when 1250 wetsuits are produced.

c) One possible reason the profit function might take this shape is labor costs. If McNeil wants to make a very large number of wetsuits in a short period of time, then that may require paying overtime for workers, and this could reduce the profit margin.

Exercises

1. Find the solution set of the following inequalities without using a calculator. Display the solution set on the number line.
   (a) $x^2 + 2x - 3 \leq 0$
   (b) $3x^2 - 7x + 2 > 0$
   (c) $-6x^2 - 13x + 5 \geq 0$
   (d) $\frac{5x-1}{x-2} > 0$
   (e) $\frac{1-x}{x} < 1$
2. Use a calculator to solve the following inequalities. Round your answer three places after the decimal.
   (a) $-9.8t^2 + 357.6t \geq 0$
   (b) $x^3 - 5x + 7 \leq -4x^2 + 18$
   (c) $\frac{x^2-2x}{x-5} > x^2 - 25$
3. The total resistance of two electronics components wired in parallel is given by $\frac{R_1 R_2}{R_1+R_2}$ where R_1 and R_2 are the individual resistances (in Ohms) of the two components.
   (a) If the resistance of R_1 is 20 Ohms, what is the maximum resistance of R_2 if the total resistance must be less than 15 Ohms?

(b) What is the maximum theoretical resistance of this circuit? How do you know?

4. A rectangular lot of land has a length that is 7 meters more than twice its width. If the area of the lot is greater than 60 square meters, what are the possible values of the widths of the lot?

Answers

1. (a) $x \in [-3, 1]$
   (b) $x \in \left(-\infty, \frac{1}{3}\right) \cup (2, +\infty)$
   (c) $x \in \left[-\frac{5}{2}, \frac{1}{3}\right]$
   (d) $x \in \left(-\infty, \frac{1}{5}\right) \cup (2, +\infty)$
   (e) $x \in (-\infty, 0) \cup \left(\frac{1}{2}, +\infty\right)$
2. (a) $t \in [0, 36.490]$
   (b) $x \in (-\infty, -4.567] \cup [-1.294, 1.861]$
   (c) $x \in (-4.667, 4.044) \cup (5, 6.623)$
3. (a) 60 Ohms
   (b) The total resistance will always be less than 20 because $y = 20$ is the horizontal asymptote of $y = \frac{20R_2}{20+R_2}$.
4. Width must be greater than 4 meters, $w > 4$.

3.6 Finding Real Zeros of Polynomial Functions

Learning Objectives

- Use polynomial long division and synthetic division to find roots
- Apply synthetic division to find the rational zeros of an unfactored polynomial (of degree >2)

Synthetic Division

As you know from years of algebra, any linear equation $ax + b = c$, can be easily solved. You have also learned how to find a solution to a quadratic equation using the quadratic formula. In this section, we will learn how to solve polynomial equations of higher degrees. We begin with a process called the *synthetic division*. It is a special case of the *polynomial long division* method. To develop the method, we will present the following example in which the divisor has the form $x - c$. We will first use the method of long division and then redo the division by the synthetic division method.

Keep in mind that if we wish to divide

$$\frac{f(x)}{D(x)}$$

where $f(x)$ is the *dividend* and $D(x)$ is the *divisor*, then

$$\frac{f(x)}{D(x)} = Q(x) + \frac{R(x)}{D(x)}$$

where $Q(x)$ is the *quotient* and $R(x)$ is the *remainder*. We can also rewrite this equation in the following form

$$f(x) = D(x)Q(x) + R(x)$$

as a solution.

Let's say we are interested in dividing $f(x) = x^3 + x^2 - 10x + 13$ by $D(x) = x - 2$.

By long division, we get

$$\begin{array}{rl}
 & x^2 + 3x - 4 \qquad \leftarrow \text{Quotient} \\
\text{Divisor} \rightarrow x - 2 \;\big) & \overline{x^3 + x^2 - 10x + 13} \qquad \leftarrow \text{Dividend} \\
 & x^3 - 2x^2 \\
 & \quad \searrow \\
 & \qquad 3x^2 - 10x \\
 & \qquad 3x^2 - 6x \\
 & \qquad\quad \searrow \\
 & \qquad\quad -4x + 13 \\
 & \qquad\quad -4x + 8 \\
 & \qquad\qquad \searrow \\
 & \qquad\qquad\quad +5 \qquad \leftarrow \text{Remainder}
\end{array}$$

As you can see, all the work involved is to manipulate the coefficients of the variables and the constants. Thus we could just as easily complete the division by omitting the variables, as long as we write the coefficients in the proper places. So the division problem will look like this:

$$\begin{array}{rl}
 & 1 + 3 - 10 \qquad \leftarrow \text{Quotient} \\
1 - 2 \;\big) & \overline{1 + 1 - 10 + 13} \qquad \leftarrow \text{Dividend} \\
 & 1 - 2 \\
 & \quad \searrow \\
 & \qquad + 3 - 10 \\
 & \qquad\quad 3 - 6 \\
 & \qquad\quad\quad \searrow \\
 & \qquad\quad\quad - 4 + 13 \\
 & \qquad\quad\quad - 4 + 8 \\
 & \qquad\quad\qquad \searrow \\
 & \qquad\quad\qquad\quad + 5 \qquad \leftarrow \text{Remainder}
\end{array}$$

Since the underlined numerals are repetitions of those immediately above them, we can shorten the process by simply deleting them. Further, since these underlined numbers are products of the numbers in the quotient by the 1 in the divisor, we eliminate this 1. Thus we get the following

$$
\begin{array}{rll}
 & 1+3-4 & \leftarrow \text{Quotient} \\
-2\,\big) & \overline{1+1-10+13} & \leftarrow \text{Dividend} \\
 & -2 & \\
 & \searrow & \\
 & +3-10 & \\
 & \quad -6 & \\
 & \quad\ \searrow & \\
 & \quad -4+13 & \\
 & \quad\ \ +8 & \\
 & \quad\ \ \searrow & \\
 & \quad\ \ +5 & \leftarrow \text{Remainder}
\end{array}
$$

It is also unnecessary to bring down the -10 and 13:

$$
\begin{array}{rll}
 & 1+3-4 & \leftarrow \text{Quotient} \\
-2\,\big) & \overline{1+1-10+13} & \leftarrow \text{Dividend} \\
 & -2 & \\
 & \searrow & \\
 & +3 & \\
 & \quad -6 & \\
 & \quad\ \searrow & \\
 & \quad -4 & \\
 & \quad\ \ +8 & \\
 & \quad\ \ \searrow & \\
 & \quad\ \ +5 & \leftarrow \text{Remainder}
\end{array}
$$

Moving the numerals upward, we get

$$
\begin{array}{rll}
 & 1+3-4 & \leftarrow \text{Quotient} \\
-2\,\big) & \overline{1+1-10+13} & \leftarrow \text{Dividend} \\
 & \underline{-2\ \ -6\ \ +8} & \\
 & +3-4\ \ +5 & \leftarrow \text{Remainder}
\end{array}
$$

When the numeral 1 (the first number in the quotient) is brought down to the last line, it will contain the remainder and the quotient,

$$
\begin{array}{rll}
-2\,\big) & \overline{1+1-10+13} & \leftarrow \text{Dividend} \\
 & \underline{-2\ \ -6\ \ +8} & \\
\text{Quotient} \rightarrow & 1+3\ \ -4\ \ +5 & \leftarrow \text{Remainder} = +5
\end{array}
$$

Therefore, the quotient is $Q(x) = x^2 + 3x - 4$ and the remainder is $R(x) = 5$. This is the method of *the synthetic division.*

Example 1

Use synthetic division to find the quotient and the remainder of

$$\frac{3x^3 - 8x + 1}{x + 2}$$

Solution

First we write the divisor $x + 2$ in the form $x - c$, as $x - (-2)$. Then use -2 as a "divisor" in the synthetic division as follows:

$$\begin{array}{r|rrrr} 2 & 3 & 0 & -8 & 1 \\ & \downarrow & -6 & 12 & -8 \\ \hline & 3 & -6 & 4 & -7 \end{array}$$

Notice that 0 is used as the coefficient of the "missing" x^2 term. Also, we wrote the coefficients of the dividend in descending order. The process of synthetic division is as follows: Bring down the first coefficient 3 and multiply by -2 (the divisor) to get -6, and then add 0 to -6 to get -6. Next, multiply -6 by -2 (the divisor) to get 12, and then add -8 to 12 to get 4. Finally, multiply 4 by -2 to get -8, and then add 1 to -8 to get the remainder, -7. As a result of this process, the quotient is

$$Q(x) = 3x^2 - 6x + 4$$

and the remainder is

$$R(x) = -7$$

In other words, since

$$\begin{aligned} f(x) &= D(x)Q(x) + R(x) \\ &= (x + 2)(3x^2 - 6x + 4) + (-7) \end{aligned}$$

Remember, this method will only work when the divisor is in the form $x - c$, that is, when the coefficient of x in the divisor is 1.

The Remainder Theorem and The Factor Theorem

This example leads us to the Remainder Theorem.

The Remainder Theorem

If a polynomial $f(x)$ of degree $n > 0$ is divided by $x - c$, then the remainder R is a constant and it is equal to the value of the polynomial when c is substituted for x. That is

$$f(c) = R$$

Proof

The division algorithm states

$$f(x) = (x - c)Q(x) + R(x)$$

since the degree of the remainder must be less than the degree of the division, $x-c$, then $R(x)$ must be a constant. We will denote it by R. In addition, if $x=c$, we will find,

$$\begin{aligned} f(c) &= (c-c)Q(x)+R \\ &= 0 \cdot Q(x)+R \\ &= R \end{aligned}$$

The Factor Theorem

If $f(x)$ is a polynomial of degree $n>0$ and $f(c)=0$, then $x-c$ is a *factor* of the polynomial $f(x)$. Further, if $x-c$ is a factor, then c is a *zero* of f.

Proof

If $f(c)=0$, then by the Remainder Theorem,

$$f(c)=R=0$$

and the division algorithm becomes,

$$\begin{aligned} f(x) &= (x-c)Q(x)+R(x) \\ &= (x-c)Q(x) \end{aligned}$$

This shows that $(x-c)$ is a factor of $f(x)$.

In addition, since $x-c$ is a factor of $f(x)$, then

$$\begin{aligned} f(x) &= (x-c)Q(x) \\ f(c) &= (c-c)Q(c) \\ &= 0 \end{aligned}$$

and c is thus a zero of f.

Example 2

Use synthetic division and the remainder and factor theorems to find the quotient $Q(x)$ and the remainder R if $f(x)=2x^3-3x^2+6$ is divided by $x-5$.

Solution

$$\begin{array}{r|rrrr} 5 & 2 & -3 & 0 & 6 \\ & \downarrow & 10 & 35 & 175 \\ \hline & 2 & 7 & 35 & 181 \end{array}$$

Hence

$$2x^3-3x^2+6=(2x^2+7x+35)(x-5)+181$$

Notice that the remainder is 181 and it can also be obtained if we simply substituted $x=5$ into $f(x)$,

$$\begin{aligned} f(5) &= 2(5)^3-3(5)^2+6 \\ &= 250-75+6 \\ &= 181 \end{aligned}$$

Example 3

If $h(x)=x^3-2x^2+5x-3$, evaluate

- $h(1)$
- $h(-2)$
- $h\left(\frac{1}{2}\right)$

Solution

We can simply use the synthetic division to evaluate $h(x)$ at the given values. By the remainder theorem, the remainder is equal to $h(c)$.

- Using synthetic division,

$$\begin{array}{r|rrrr} 1 & 1 & -2 & 5 & -3 \\ & \downarrow & 1 & -1 & 4 \\ \hline & 1 & -1 & 4 & 1 \end{array}$$

Hence, $h(1) = 1$.

- By synthetic division,

$$\begin{array}{r|rrrr} -2 & 1 & -2 & 5 & -3 \\ & \downarrow & 2 & 8 & 26 \\ \hline & 1 & -4 & 13 & -29 \end{array}$$

Hence, $h(-2) = -29$.

- By synthetic division,

$$\begin{array}{r|rrrr} \frac{1}{2} & 1 & -2 & 5 & -3 \\ & \downarrow & \frac{1}{2} & \frac{-3}{4} & \frac{17}{8} \\ \hline & 1 & \frac{-3}{2} & \frac{17}{4} & \frac{-7}{8} \end{array}$$

Hence, $h\left(\frac{1}{2}\right) = \frac{-7}{8}$.

Example 4

Show that $x + 3$ is a factor of $g(x) = x^4 + 2x^3 - 3x^2 + 4x + 12$. Find the quotient $Q(x)$ and express $f(x)$ is factored form.

Solution

By the factor theorem, if $f(c) = 0$, then $x - c$ is a factor of the polynomial. In other words, if the synthetic division produces a remainder equal to zero, then c is a factor of the polynomial. Using the synthetic division with $c = -3$,

$$\begin{array}{r|rrrr} -3 & 1 & 2 & -3 & 4 & 12 \\ & \downarrow & -3 & 3 & 0 & -12 \\ \hline & 1 & -1 & 0 & 4 & 0 \end{array}$$

Hence, $g(-3) = 0$, and the quotient is

$$Q(x) = x^3 - x^2 + 4$$

so that $g(x)$ can be written as

$$g(x) = (x - (-3))(x^3 - x^2 + 4)$$
$$g(x) = (x + 3)(x^3 - x^2 + 4)$$

Rational Zero Theorem

Recall that the zeros of a polynomial represent the $x-$intercepts on the graph. It is easy to find the zeros of first or second degree polynomials but as the degree increases, the problem of determining the zeros becomes more difficult. There is a general method however for finding the zeros of polynomials (if they exist) for degrees $n > 2$. Before we state the method it is important to recall one of the fundamental properties of integers: *Every positive integer can be expressed as a product of prime numbers.* For example, $25 = 5^2, 22 = 2 \cdot 11$, and so on. Here 5, 2, and 11 are prime numbers.

The Rational Zero Theorem

Given the polynomial

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

$a_n \neq 0$ and n is a positive integer. If the coefficients are integers and $\frac{p}{q}$ is a rational zero in lowest terms, then p is a divisor of a_0 and q is a divisor of a_n.

Example 5

Use the theorem above and synthetic division to find all the possible rational zeros of the polynomial

$$f(x) = x^3 - 2x^2 - x + 2$$

Solution

From the rational zero theorem, $\frac{p}{q}$ is a rational zero of the polynomial f. So p is a divisor of 2 and q is a divisor of 1. Hence, p can take the following values: -1, 1, -2, 2 and q can be either -1 or 1. Therefore, the possible values of $\frac{p}{q}$ are

$$\frac{p}{q} : -1, 1, -2, 2$$

So there are four possible zeros. Of these four, not more than three can be zeros of f because f is a polynomial with degree 3. To test which of the four possible candidates are zeros of f, we use the synthetic division. Recall from the remainder theorem if $f(c) = 0$, then c is a zero of f. We have

$$\begin{array}{r|rrrr} 2 & 1 & -2 & -1 & 2 \\ & \downarrow & 2 & 0 & -2 \\ \hline & 1 & 0 & -1 & 0 \end{array}$$

Hence, 2 is a zero of f. Further, by the division algorithm,

$$f(x) = (x-c)Q(x) + R(x)$$
$$= (x-2)(x^2-1) + 0$$

The remaining zeros of f are simply the zeros of $Q(x) = x^2 - 1$ which is easier to manipulate,

$$Q(x) = x^2 - 1$$
$$= (x-1)(x+1)$$

and thus the remaining zeros are -1 and 1. Thus the rational zeros of f are -1, 1, and 2.

Example 6

Use the theorem above and synthetic division to find all the possible rational zeros of the polynomial

$$f(x) = x^3 - 2x^2 - 5x + 6$$

Solution

Assume $\frac{p}{q}$ is a rational zero of f. By the rational zero theorem, p is a divisor of 6 and q is a divisor of 1. Thus p and q can assume the following respective values

$$p : 1, -1, 2, -2, 3, -3, 6, -6$$

and

$$q : -1, 1$$

Therefore, the possible rational zeros will be

$$\frac{p}{q} : -1, 1, -2, 2, -3, 3, -6, 6$$

Notice that with these choices for p and q there could be $8 \cdot 2 = 16$ rational zeros. But, eight of them are duplicates. For example $\frac{1}{-1} = \frac{-1}{1} = -1$. The next step is to test all these values by the synthetic division (we'll let you do this on your own for practice) and we finally find that

$$1, -2, \text{ and } 3$$

are zeros of f. That is

$$f(x) = x^3 - 2x^2 - 5x + 6$$
$$= (x-1)(x+2)(x-3)$$

Keep in mind that the rational zero theorem has only a limited ability in finding the zeros of a polynomial equation. For example, the equation $x^2 - 3x + 1 = 0$ can not be solved by using the rational zero theorem because the roots of this equation are *irrational* numbers.

Graphing Polynomials Using Rational Zero Theorem

Example 7

Graph the polynomial function $h(x) = 2x^3 - 9x^2 + 12x - 5$.

Solution

Notice that the leading term is $2x^3$, where $n = 3$ odd and $a_n = 2 > 0$. This tells us that the end behavior will take the shape of a power function with an odd exponent.

Here, as you can see, there is no straight-forward way to find the zeros of $h(x)$. However, with the use of the factor theorem and the synthetic division, we can find the rational roots of $h(x)$. First, we use the rational zero theorem and find that the possible rational zeros are

$$\frac{p}{q} : -1, 1, -2, 2, -5, 5, -\frac{5}{2}, \frac{5}{2}$$

testing all these numbers by the synthetic division,

$$\begin{array}{r|rrrr} -1 & 2 & -9 & 12 & -5 \\ & \downarrow & -2 & 11 & -23 \\ \hline & 2 & -11 & 23 & -28 \end{array}$$

-1 is not a root. Now let's test $x = 1$.

$$\begin{array}{r|rrrr} 1 & 2 & -9 & 12 & -5 \\ & \downarrow & -2 & -7 & 5 \\ \hline & 2 & -7 & 5 & -28 \end{array}$$

we find that 1 is a zero of h and so we can re-write $h(x)$,

$$h(x) = (2x^2 - 7x + 5)(x - 1)$$

Looking at quadratic part,

$$2x^2 - 7x + 5 = (2x - 5)(x - 1)$$

and so

$$h(x) = (2x - 5)(x - 1)^2$$

Thus 1 and $\frac{5}{2}$ are the $x-$intercepts of $h(x)$. The $y-$intercept is

$$h(0) = -5$$

Further, the synthetic division can be also used to form a table of values for the graph of $h(x)$:

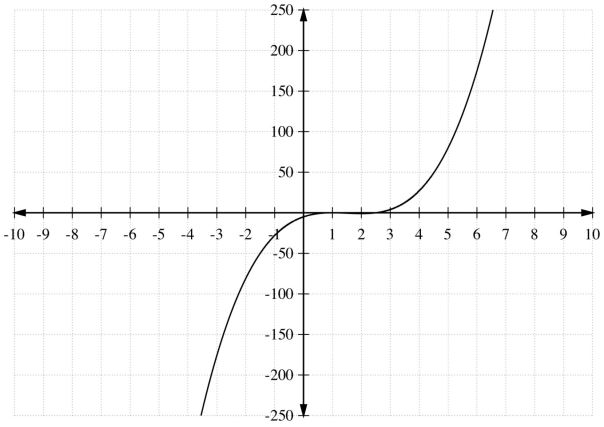
x	-1	0	1	2	$\frac{5}{2}$	3
$h(x)$	-28	-5	0	-1	0	4

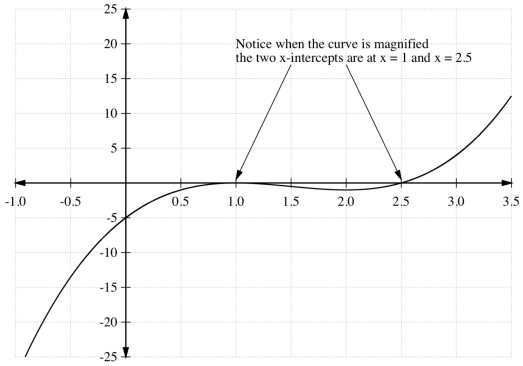
We choose test points from each interval and find $g(x)$.

Table 3.13:

Interval	**Test Value** x	$h(x)$	**Sign of** $h(x)$	**Location of points on the graph**
$(-\infty, 1)$	-1	-28	-	below the $x-$axis
$\left(1, \frac{5}{2}\right)$	$\frac{3}{2}$	$\frac{-11}{4}$	-	below the $x-$axis
$\left(\frac{5}{2}, \infty\right)$	3	4	$+$	above the $x-$axis

From this information, the graph of $h(x)$ is shown in the two graphs below. Notice that the second graph is a magnification of $h(x)$ in the vicinity of the $x-$axis.





Descartes' Rule of Signs

One challenge for finding the zeros of polynomials is that you may not know how many roots a polynomial actually has. Descartes Rule of Signs helps with this problem. Descartes' Rule of Signs is given as a procedure for finding the number of possible positive and negative roots of a polynomial.

Descartes Rule of Signs

Given any polynomial, $p(x)$,

1. Write it with the terms in descending order, i.e. from the highest degree term to the lowest degree

term.
2. Count the number of sign changes of the terms in $p(x)$. Call the number of sign changes n.
3. Then the number of *positive* roots of $p(x)$ is less than or equal to n.
4. Further, the possible number of positive roots is $n, n-2, n-4, \ldots$
5. To find the number of *negative* roots of $p(x)$, write $p(-x)$ in descending order as above (i.e. change the sign of all terms in $p(x)$ with odd powers), and repeat the process above. Then the *maximum* number of negative roots is n.

Example 8

Use Descartes Rule of Signs to identify the possible number of positive and negative roots of $f(x) = -2x^3 + x^2 - 3x^5 + 5x - 1$.

Solution

First, re-write $f(x)$ in descending order

$$f(x) = -3x^5 - 2x^3 + x^2 + 5x - 1.$$

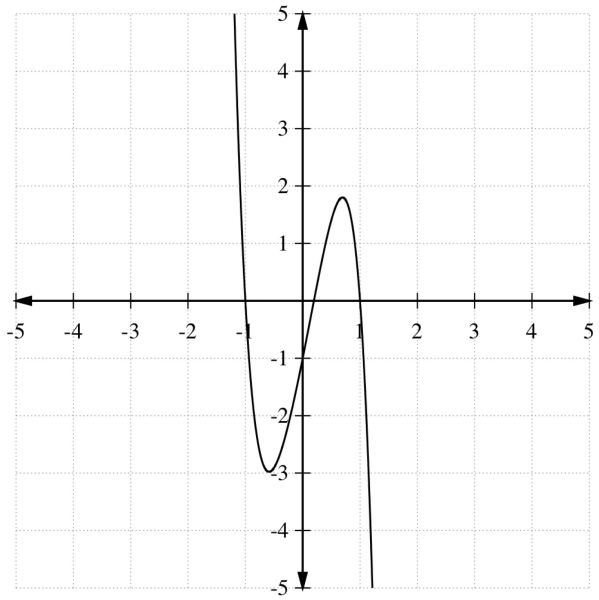
The number of sign changes of $f(x)$ is 2, so the number of positive roots is either 2 or 0.

For the negative roots, write

$$f(-x) = -3x^5 + 2x^3 + x^2 - 5x - 1$$

The number of sign changes of $f(-x)$ is 2, so the maximum number of negative roots is 2.

The graph of $f(x)$ below shows that there is one negative root and two positive roots.



Exercises

1. Use (i) long division and (ii) synthetic division to perform the divisions. Express each result in the form $f(x) = D(x) \cdot Q(x) + R$.
   (a) $5x^5 - 3x^4 + 2x^3 + x^2 - 7x + 3$ by $x - 2$

(b) $-4x^6 - 5x^3 + 3x^2 + x + 7$ by $x - 1$
(c) $2x^3 - 5x^2 + 5x + 11$ by $x - \frac{1}{2}$

2. Use synthetic division to find $Q(x)$ and $f(c)$ so that $f(x) = (x - c)Q(x) + f(c)$ if $f(x) = -3x^4 - 3x^3 + 3x^2 + 2x - 4$ and $c = -2$
3. If $f(x) = x^3 + 2x^2 - 10x + 10$, use synthetic division to determine the followings:

   (a) $f(-1)$
   (b) $f(-3)$
   (c) $f(0)$
   (d) $f(4)$
   (e) What are the factors of $f(x)$?

4. Find k so that $x - 2$ is a factor of $f(x) = 3x^3 + 4x^2 + kx - 20$
5. Use synthetic division to determine all the zeros of the polynomials.

   (a) $f(x) = 3x^3 - 7x^2 + 8x - 2$
   (b) $g(x) = 4x^4 - 4x^3 - 7x^2 + 4x + 3$

6. Graph the polynomial function $f(x) = x^3 - 2x^2 - 5x + 6$ by using synthetic division to find the x–intercepts and locate the y–intercepts.
7. Graph the polynomial function $h(x) = x^3 - 3x^2 + 4$ by using synthetic division to find the x–intercepts and locate the y–intercepts.
8. Let $f(x) = 2x^3 - 5x^2 - 4x + 3$. Find the solution of the following equations:

   (a) $f(x) = 0$
   (b) $f(2x) = 0$

9. Refer back to Problem 6 above. Use the graph of $f(x)$ to find the solution set of the inequality $x^3 - 2x^2 - 5x + 6 \leq 0$.
10. Use the graph of $f(x) = x(x - 1)(x + 2)$ to find the solution set of the inequality $x(x - 1)(x + 2) > 0$.

Answers

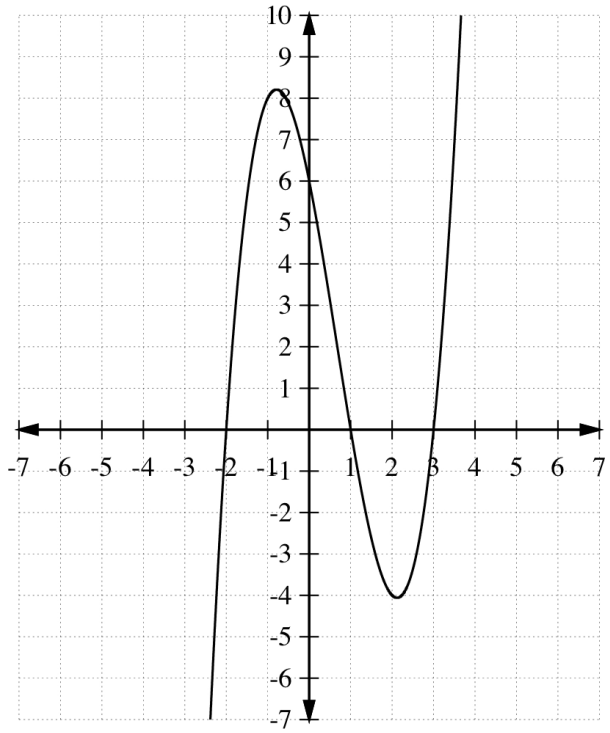
1. Answers:

   (a) $f(x) = (x - 2)(5x^4 + 7x^3 + 16x^2 + 33x + 59) + 121$
   (b) $f(x) = (x - 1)(-4x^5 - 4x^4 - 4x^3 - 9x^2 - 6x - 5) + 2$
   (c) $f(x) = (x - \frac{1}{2})(2x^2 - 4x + 3) + \frac{25}{2}$

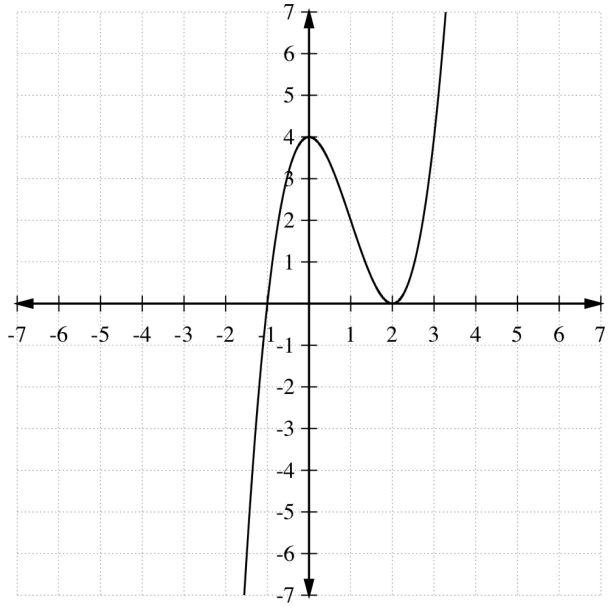
2. $Q(x) = -3x^3 + 3x^2 - 3x + 8, f(-2) = -20$
3. (a) 21
   (b) 31
   (c) 10
   (d) 66
4. $k = -10$
5. Answers:

   (a) $\frac{1}{3}$
   (b) $-1, -\frac{1}{2}, 1, \frac{3}{2}$

6. x–intercepts: -2, 1, 3; y–intercepts: 6



7. $x-$intercepts: -1, 2; $y-$intercepts: 4



8. Answers:

   (a) $-1, \frac{1}{2}, 3$
   (b) $-\frac{1}{2}, \frac{1}{4}, \frac{3}{2}$

9. $x \in (-\infty, -2] \cup [1, 3]$
10. $x \in (-2, 0) \cup (1, +\infty)$

3.7 Approximating Real Zeros of Polynomial Functions

Learning Objectives

- Understand the statement of the Intermediate Value Theorem
- Apply the intermediate value theorem to find bounds on the zeros of a function
- Use numerical methods to find roots of a polynomial

Intermediate Value Theorem & Bounds on Zeros

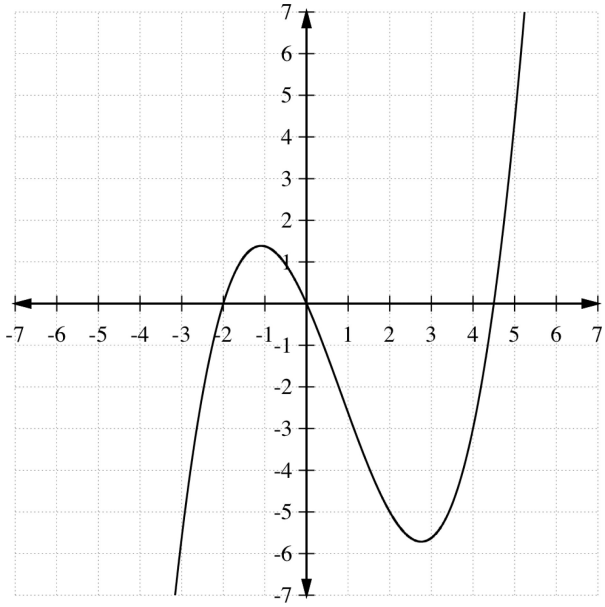
The intermediate value theorem offers one way to find roots of a continuous function. Recall that our informal definition of *continuous* is that a function is continuous over a certain interval if it has no breaks, jumps asymptotes, or holes in that interval. Polynomial functions are continuous for all real numbers x. Rational functions are often not continuous over the set of real numbers because of asymptotes or holes in the graph. But for intervals without holes, rational functions are continuous.

If we know a function is continuous over some interval $[a, b]$, then we can use the intermediate value theorem:

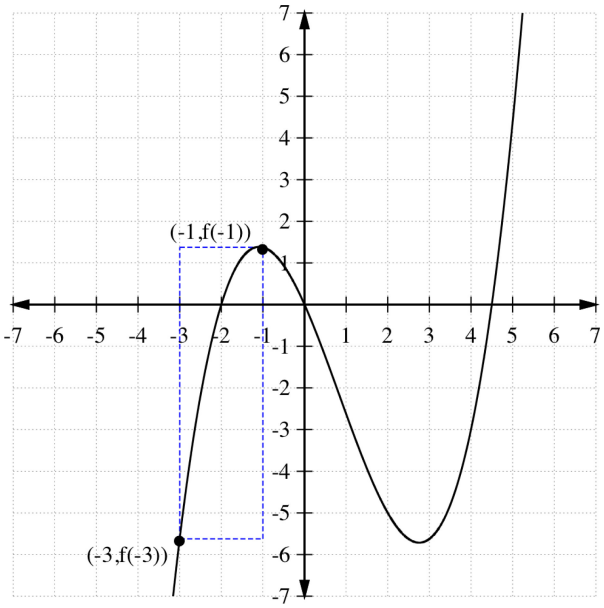
Intermediate Value Theorem

If $f(x)$ is continuous on some interval $[a, b]$ and n is between $f(a)$ and $f(b)$, then there is some $c \in [a, b]$ such that $f(c) = n$.

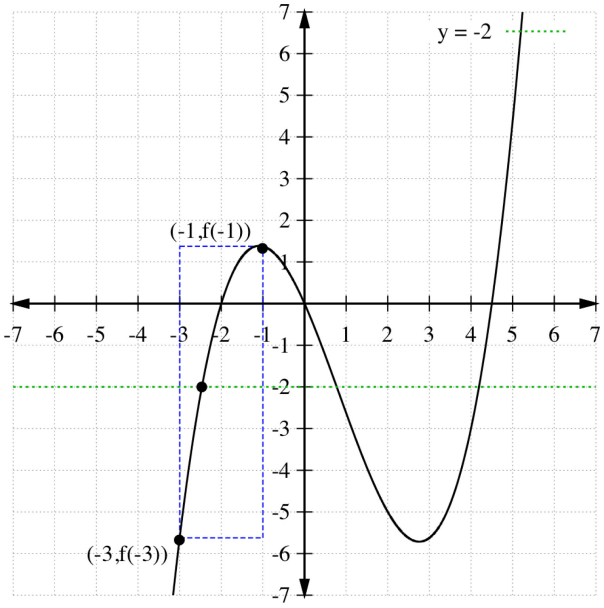
The following graphs highlight how the intermediate value theorem works. Consider the graph of the function $f(x) = \frac{1}{4}\left(x^3 - \frac{5x^2}{2} - 9x\right)$ below on the interval [-3, -1].



$f(-3) = -5.625$ and $f(-1) = 1.375$. If we draw bounds on [-3, -1] and $[f(-3), f(-1)]$, then we see that for any $y-$value between $y = -5.625$ and $y = 1.375$, there must be an x value in [-3, -1] such that $f(x) = y$.



So, for example, if we choose $c = -2$, we know that for some $x \in [-3, -1], f(x) = -2$, even though solving this by hand would be a chore!



The Bounds on Zeros Theorem

The Bounds on Zeros Theorem is a corollary to the Intermediate Value Theorem:

Bounds on Zeros Theorem

If f is continuous on $[a, b]$ and there is a sign change between $f(a)$ and $f(b)$ (that is, $f(a)$ is positive and $f(b)$ is negative, or vice versa), then there is a $c \in (a, b)$ such that $f(c) = 0$.

The bounds on zeros theorem is a corollary to the intermediate value theorem because it is not fundamentally different from the general statement of the IVT, just a special case where $n = 0$.

Looking back at $f(x) = \frac{1}{4}\left(x^3 - \frac{5x^2}{2} - 9x\right)$ above, because $f(-3) < 0$ and $f(-1) > 0$, we know that for some

$x \in [-3, -1], f(x)$ has a root. In fact, that root is at $x = 2.$ and we can test that using synthetic division or by evaluating $f(2)$ directly.

Example 1

Show that $f(x) = -3x^3 + 5x$ has at least one root in the interval [1, 2]

Solution

Since $f(x)$ is a polynomial we know that it is continuous. $f(1) = 2$ and $f(2) = -14$. Let $n = 0 \in [-14, 2]$. Applying the Intermediate Value Theorem, there must exist some point $c \in [1, 2]$ such that $f(c) = 0$. This proves that $f(x)$ has a root in [1, 2].

Example 2

The table below shows several sample values of a polynomial $p(x)$.

x	-4	-2	0	1	4	6	8	10	15	18
$p(x)$	44.15	6.62	-4.12	-4.09	1.16	0	-8.74	-24.07	-49.89	3.41

Based on the information in the table

(a) What is the minimum number of roots of $p(x)$?

(b) What are bounds on the roots of $p(x)$ that you identified in (a)?

Solution

Since $p(x)$ is a polynomial we already know that it is continuous. We can use the Intermediate Value Theorem to identify roots by looking at when $p(x)$ changes from negative to positive, or from positive to negative.

(a) There are four sign changes of $p(x)$ in the table, so at minimum, $p(x)$ has four roots.

(b) The roots are in the following intervals $x \in [-2, 0], x \in [1, 4], x \in [15, 18]$, and the table also tells us that one root is at $x = 6$.

Approximate Zeros of Polynomials Functions

In calculus you will learn several methods for numerically approximating the roots of functions. In this section we show one elementary numerical method for finding the zeros of a polynomial which takes advantage of the Intermediate Value Theorem.

Given a continuous function $g(x)$,

1. Find two points such that $g(a) > 0$ and $g(b) < 0$. Once you have found these two points, you can iteratively use the steps below to find the root of $g(x)$ on the interval $[a, b]$. (Note, we will assume $a < b$, the same algorithm works with minor adjustments if $b > a$)
2. Evaluate $g\left(\frac{a+b}{2}\right)$.
   (a) If $g\left(\frac{a+b}{2}\right) = 0$, then the root is $x = \frac{a+b}{2}$.
   (b) If $g\left(\frac{a+b}{2}\right) > 0$, replace a with $\frac{a+b}{2}$. and repeat steps 1-2 using $\left[\frac{a+b}{2}, b\right]$
   (c) If $g\left(\frac{a+b}{2}\right) < 0$, replace b with $\frac{a+b}{2}$. and repeat steps 1-2 using $\left[a, \frac{a+b}{2}\right]$

This algorithm will not usually find the exact root of $g(x)$, but it will allow you to find a reasonably small interval for the root. For example, you could repeat this process enough times so that you find an interval with $|a - b| < 0.01$, and you will know the root of $g(x)$ within a reasonably good approximation. The quality

of the approximation you use (and the number of steps you use) will depend on why you are looking for the root. For most applications coming within 0.01 of the root is a reasonable approximation, but for some applications (such as building a bridge or launching a rocket) you need much more accuracy.

Example 3

Show the first 5 iterations of finding the root of $h(x) = x^2 - x - 1$ using the starting values $a = 0$ and $b = 2$.

Solution:

1. First we verify that there is a root between $x = 0$ and $x = 2$. $h(0) = -1$ and $h(2) = 1$ so we know there is a root in the interval [0, 2]. Check $h\left(\frac{2+0}{2}\right) = h(1) = -1$. Since $-1 < 0$ we know the root is between $x = 1$ and $x = 2$, and we use the new interval [1, 2].
2. Now we use the interval [1, 2]. $h\left(\frac{1+2}{2}\right) = h(1.5) = -0.25$. Since $-0.25 < 0$, we use the interval [1.5, 2].
3. $h\left(\frac{1.5+2}{2}\right) = h(1.75) = 0.31$. Since $0.31 > 0$, we know that the zero is in the interval [1.5, 1.75].
4. $h\left(\frac{1.5+1.75}{2}\right) = h(1.625) = 0.02$. Since $0.02 > 0$, we know the root is between 1.5 and 1.625.
5. $h\left(\frac{1.5+1.625}{2}\right) = h(1.5620) = -0.12$. Since $-0.12 < 0$, we know the root is between 1.5620 and 1.625.

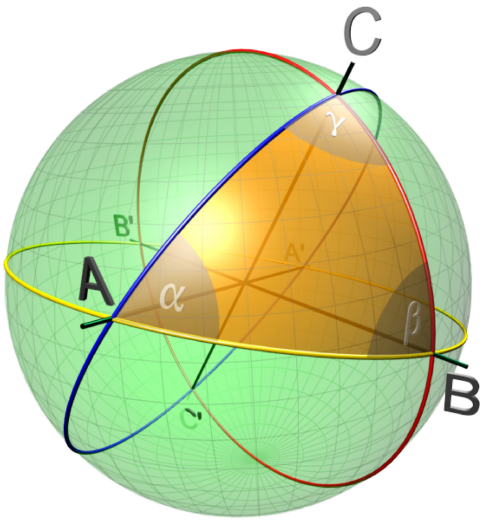
This example shows that after five iterations we have narrowed the possible location of the root to within 0.06 units. Not bad!

Recall that we have already reviewed using the **CALC** menu on a graphing calculator to find the roots of a function. This algorithm is not the one used by a calculator, but the calculator uses a similar, more efficient, algorithm for approximating the root of a function to 13 decimal places. When the calculator prompts for a **GUESS**? it is asking for a starting value to run the iterations.

Optional: An Interesting Corollary of the IVT

One surprising result of the Intermediate Value Theorem is that if you draw any great circle around the globe, then there must two antipodal points on that great circle that have exactly the same temperature.

Recall that a *great circle* is a path around a sphere that gives the shortest distance between any two points on the sphere. The equator is a great circle around the globe. *Antipodal* points are two points on opposite sides of the sphere. In the diagram below, B and B' are antipodal.



(*Source:* http://commons.wikimedia.org/wiki/File:Spherical_triangle_3d_opti.png ***License:*** GNU FDL)

For an informal proof of this result, look at the the image of a sphere with three great circles above. Suppose that the temperature at B is 75° and the temperature B' is 50°. The difference between the temperature at B and at B' is $75 - 50 = 25$. Now imagine rotating the segment $\overline{BB'}$ around the blue great circle. When the segment has rotated 180 degrees (i.e. when B has rotated to where B' is), then The difference between the temperatures at these two points is $50 - 75 = -25$. Since temperatures vary continuously, by the intermediate value theorem, there must be some point on that circle when the difference was 0, implying two antipodal points had the same temperature.

Notice that this little demonstration does not tell us which two antipodal points had the same temperature, only that there must be two such points on any great circle.

Exercises

For Exercises 1-5 use the intermediate value theorem to show the bounds on the zeros of each function. Your bounds should be within the whole number

1. $f(x) = 2x^3 - 3x + 4$
2. $g(x) = -5x^2 + 8x + 12$
3. $h(x) = \frac{1}{2}x^4 - x^3 - 3x^2 + 1$
4. $j(x) = -\frac{2}{x^2+1} + \frac{1}{2}$
5. $k(x)$ is a polynomial and selected values of $k(x)$ are given in the following table:

x	-3	-2	-1	0	1	2	3
$k(x)$	-23.5	-1	0.5	-1	$.5$	-1	-23.5

6. Stephen argues the function $r(x) = \frac{4x+1}{x+3.5}$ has two zeros based on the following table and an application of the Bounds on Zeros Theorem. What is faulty about Stephen's reasoning?

x	-5	-4	-3	-2	-1	0	1	2	3	4
$r(x)$	12.67	30	-22	-4.67	-1.20	0.29	1.11	1.64	2.0	2.27

7. Apply the numerical algorithm five times to find a bound on the zeros of the following functions given the indicated starting values. What is your final estimate for the zero?
   (a) $k(x) = x^4 - 3x + 1$ on [0, 1]
   (b) $b(x) = -0.1x^5 + 3x^3 - 5x^2$ on [1, 3]
   (c) $c(x) = \frac{3x^2-2}{x^4+2}$ on [0, 2]

Answers

1. There is a zero in [-2, -1]
2. There is a zero in [-1, 0] and [2, 3]
3. There is a zero in [-2, -1], [-1, 0], [0, 1], and [3, 4]
4. There is a zero in [-2, -1] and [1, 2]
5. There is a zero in [-2, -1], [-1, 0], [0, 1] and [1, 2]
6. Unlike the previous question which specified the function was a polynomial (and hence continuous), $r(x) = \frac{4x+1}{x+3.5}$ has a vertical asymptote at $x = -3.5$, so it is not continuous in the interval [-4, -3]. Therefore we cannot use the Bounds on Zeros theorem to claim there is a zero in that interval.
7. (a) The zero is in [0.3125, 0.34825]
   (b) The zero is in [1.5000, 1.5625]
   (c) [0.8125, 0.875]

3.8 The Fundamental Theorem of Algebra

Learning objectives

- Understand the statement of the theorem and how to apply it to various functions
- Understand the conjugate root theorem

Complex Roots of Polynomial Functions

A polynomial function with real coefficients does not necessarily have real zeros. You may recall that the quadratic equation $f(x) = ax^2 + bx + c$, where a, b, and c are real numbers has real zeros if and only if the discriminant $b^2 - 4ac \geq 0$. Otherwise, the quadratic equation has complex roots. For example, the zeros of the quadratic equation

$$f(x) = x^2 + 3x + 4$$

can be found by using the quadratic formula ((insert cross reference?)) as follows

$$\begin{aligned} x &= \frac{-3 \pm \sqrt{9 - 16}}{2} \\ &= \frac{-3 \pm \sqrt{-7}}{2} \\ &= \frac{-3 \pm \sqrt{7}i}{2} \end{aligned}$$

Therefore the zeros are $x = -\frac{3}{2} + \frac{\sqrt{7}}{2}i$ and $x = -\frac{3}{2} - \frac{\sqrt{7}}{2}i$.

Another example, the polynomial function

$$f(x) = 2x^3 - 5x^2 + 4x - 3$$

can be factored using synthetic division into

$$f(x) = (x - 3)(2x^2 + x + 1)$$

Using the quadratic formula on the second term, we find all the zeros to be,

$$3, -\frac{1}{4} + \frac{\sqrt{7}}{4}i, -\frac{1}{4} - \frac{\sqrt{7}}{4}i$$

and thus $f(x)$ can be written as

$$\begin{aligned} f(x) &= 2x^3 - 5x^2 + 4x - 3 \\ &= (x - 3)\left[x - \left(-\frac{1}{4} + \frac{\sqrt{7}}{4}i\right)\right]\left[x - \left(-\frac{1}{4} - \frac{\sqrt{7}}{4}i\right)\right] \end{aligned}$$

Written in this form, $f(x)$ is a *complex polynomial function* written in factored form.

Example 1

Write $g(x) = x^2 + x + 1$ as a complex polynomial in factored form.

Solution

Notice that $g(x)$ has no real roots. You can see this in the graph of $g(x)$, or by looking at the discriminant, $b^2 - 4ac = 1 - 4 = -3$.

<insert graph>

Using the quadratic formula, the roots of $g(x)$ are

$$\begin{aligned} x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2z} \\ &= \frac{-1 \pm \sqrt{-3}}{2} \\ &= -\frac{1}{2} + i\frac{\sqrt{3}}{2} \text{ or } -\frac{1}{2} - i\frac{\sqrt{3}}{2} \end{aligned}$$

Finally, writing $g(x)$ in factored form,

$$g(x) = \left[x - \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)\right]\left[x - \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right)\right]$$

Fundamental Theorem of Algebra

From the results above, we conclude that all polynomials are "factorable" as products of linear factors. These factors are based on the zeros of the polynomial functions. We present the following important four theorems in the theory of complex zeros of polynomial functions:

The Fundamental Theorem of Algebra

If $f(x)$ is a polynomial of degree $n \geq 1$, then $f(x)$ has at least one zero in the complex number domain. In other words, there is at least one complex number c such that $f(c) = 0$.

There is no rigorous proof for the fundamental theorem of algebra. Some mathematicians even believe that such proof may not exist. However, the theorem is considered to be one of the most important theorems in mathematics. A corollary of this important theorem is the *factorization theorem*,

Theorem 2: The Factorization Theorem

If

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where $a_n \neq 0$, and n is a positive integer, then

$$f(x) = a_n(x - c_1)(x - c_2) \cdots (x - c_0)$$

where the numbers c_i are complex numbers.

Theorem 3: The n–Roots Theorem

If $f(x)$ is a polynomial of degree n, where $n \neq 0$, then $f(x)$ has, at most, n zeros.

Notice that this theorem does not restrict that the zeros must be distinct. In other words, multiplicity of the zeros is allowed. For example, the quadratic equation $f(x) = x^2 + 6x + 9$ has one zero, -3, and we say that the function has -3 as a double zero or one zero with multiplicity $k = 2$. In general, if

$$f(x) = (x-c)^k q(x) \quad \text{and} \quad q(c) \neq 0$$

then c is a zero of the polynomial f and of multiplicity k. For example,

$$f(x) = (x-2)^3(x+5)$$

has 2 as one zero with $k = 3$ and -5 as a zero with $k = 1$.

Conjugate Pairs Theorem

Theorem: The Conjugate Root Theorem

If $f(z)$ is a polynomial of degree n, with $n \neq 0$ and with real coefficients, and if $f(z_0) = 0$, where $z_0 = a + bi$, then $f(z_0^*) = 0$. Where z_0^* is the complex conjugate of z_0.

This is a fascinating theorem! It says basically that if a complex number is a zero of a polynomial, then its complex conjugate *must* also be a zero of the same polynomial. In other words, complex roots (or zeros) exist in *conjugate pairs* for the same polynomial. For example, the polynomial function

$$f(x) = x^2 - 2x + 2$$

has two zeros: one is the complex number $1 + i$. By the conjugate root theorem, $1 - i$ is also a zero of $f(x) = x^2 - 2x + 2$. We can easily prove that by multiplication:

$$\begin{aligned}
[x-(1+i)]\,[x-(1-i)] &= (x-1-i)(x-1+i) \\
&= x^2 - x + ix - x + 1 - i - ix + i + 1 \\
&= x^2 - 2x + 2
\end{aligned}$$

Example 2

What is the form of the polynomial $f(x)$ if it has the following numbers as zeros: $\frac{-1}{3}, 1-i$ and $2i$?

Solution

Since the numbers $2i$ and $1+i$ are zeros, then they are roots of $f(x) = 0$. It follows that they must satisfy the conjugate root theorem. Thus $-2i$ and $1-i$ must also be roots to $f(x)$. Therefore,

$$f(x) = \left(x + \frac{1}{3}\right)[x-(1-i)][x-(1+i)][x-(2i)][x-(-2i)]$$

Simplifying,

$$f(x) = \left(x + \frac{1}{3}\right)(x-1+i)(x-1-i)(x-2i)(x+2i)$$

After multiplying we get,

$$f(x) = \frac{1}{3}(3x^5 - 5x^4 + 13x^3 - 19x^2 + 4x + 4)$$

which is a fifth degree polynomial. Notice that the total number of zeros is also 5.

Example 3

What is the multiplicity of the zeros to the polynomial

$$g(x) = x^4 - 6x^3 + 18x^2 - 54x + 81$$

Solution

With the help of the rational zero theorem and the synthetic division, we find that $x = 3$ is a zero of $g(x)$,

$$\begin{array}{r|rrrrr} 3 & 1 & -6 & 18 & -54 & 81 \\ & \downarrow & 3 & -9 & 27 & -81 \\ \hline & 1 & -3 & 9 & -27 & 0 \end{array}$$

$$g(x) = x^4 - 6x^3 + 18x^2 - 54x + 81 = (x-3)(x^3 - 3x^2 + 9x - 27)$$

Using synthetic division on the quotient, we find that 3 is again a zero:

$$\begin{array}{r|rrrr} 3 & 1 & -3 & 9 & -27 \\ & \downarrow & 3 & 9 & -27 \\ \hline & 1 & 0 & 9 & 0 \end{array}$$

or from the n–Root Theorem (Theorem 3), we write our solution as

$$\begin{aligned} g(x) &= (x-3)(x-3)(x^2+9) \\ &= (x-3)^2(x-3i)(x+3i) \end{aligned}$$

So 3 is a double zero ($k = 2$) and $3i$ and $-3i$ are each of $k = 1$.

Exercises

In problems 1-5, find a polynomial function with real coefficients that has the given numbers as its zeros.

1. $1, 2, i$
2. $2, 2, 1 - i$
3. $i, i, 0, 2i$
4. $1, 1, \left(1 - i\sqrt{3}\right)$
5. $0, 0, 2i$
6. If $i - 1$ is a root of the polynomial $f(x) = x^4 + 2x^3 - 4x - 4$, find all other roots of f.
7. If $-2i$ is a zero of the polynomial $f(x) = x^4 + x^3 - 2x^2 + 4x - 24$, find all other zeros of f.

In Problems 8-10, determine whether the given number is a zero of the given polynomial. If so, determine its multiplicity.

8. $f(x) = 9x^4 - 12x^3 + 13x^2 - 12x + 4, x = \frac{2}{3}$
9. $f(x) = x^4 - 4x^3 + 5x^2 - 4x + 4, x = 2$
10. $f(x) = 3x^5 - 4x^4 + 2x^3 - \frac{3}{4}x^2 + 2x + 12, x = -\frac{2}{3}$

Answers

1. $x^4 - 3x^3 + 3x^2 - 3x + 2$
2. $x^4 - 6x^3 + 14x^2 - 16x + 8$
3. $x^7 + 6x^5 + 9x^3 + 4x$
4. $x^4 - 4x^3 + 9x^2 - 10x + 4$
5. $x^4 + 4x^2$
6. $-i - 1, \sqrt{2}, -\sqrt{2}$
7. $2i, -3, 2$
8. Yes; $k = 3$
9. Yes; $k = 2$
10. No

Chapter 4

Analyzing Exponential and Logarithmic Functions

4.1 Inverse Functions

Learning objectives

- Find the inverse of a function
- Determine if a function is invertible
- State the domain and range for a function and its inverse
- Graph functions and their inverses
- Use composition to verify if two functions are inverses.

Introduction

In chapters 1 and 2 we have examined properties and applications of several families of functions. In this chapter, we will focus on two related functions: exponential functions,and logarithmic functions. These two functions have a special relationship with one another: they are **inverses** of each other. In this first lesson we will develop the idea of inverses, both algebraically and graphically, as background for studying these two types of functions in depth. We will begin with a familiar, every-day example of two functions that are inverses.

Functions and inverses

In the United States, we measure temperature using the Fahrenheit scale. In other countries, people use the Celsius scale. The equation $C = 5/9$ (F - 32) can be used to find C, the Celsius temperature, given F, the Fahrenheit temperature. If we write this equation using function notation, we have $t(x) = \frac{5}{9}(x - 32)$. The input of the function is a Fahrenheit temperature, and the output is a Celsius temperature. For example, the freezing point on the Fahrenheit scale is 32 degrees. We can find the corresponding Celsius temperature using the function:

$$t(32) = \frac{5}{9}(32 - 32) = \frac{5}{9} \cdot 0 = 0$$

This function allows us to convert a Fahrenheit temperature into Celsius, but what if we want to convert from Celsius to Fahrenheit?

Consider again the equation above: $C = \frac{5}{9}(F - 32)$. We can solve this equation to isolate F:

Table 4.1:

$C = \frac{5}{9}(F - 32)$
$\frac{9}{5}C = \frac{9}{5} \times \frac{5}{9}(F - 32)$ $\frac{9}{5}C = F - 32$ $\frac{9}{5}C + 32 = F$

If we write this equation using function notation, we get $f(x) = \frac{9}{5}x + 32$. For this function, the input is the Celsius temperature, and the output is the Fahrenheit temperature. For example, if $x = 0, f(0) = \frac{9}{5}(0) + 32 = 0 + 32 = 32$.

Now consider the functions $t(x) = \frac{5}{9}(x - 32)$ and $f(x) = \frac{9}{5}x + 32$ together. The input of one function is the output of the other. This is an informal way of saying that these functions are **inverses.** Formally, the inverse of a function is defined as follows:

Table 4.2:

Inverse functions

Functions *f(*x*) and* g*(*x*) are said to inverses if*

$f(g(x)) = g(f(x)) = x$

Or, using the composite function notation of Chapter 1:

$f \circ g = g \circ f = x$

The following notation is used to indicate inverse functions:
If $f(x)$ and $g(x)$ are inverse functions, then

$f(x) = g^{-1}(x)$ and $g(x) = f^{-1}(x)$

The following notation is also used: $f = g^{-1}$ and $g = f^{-1}$.

Note that $f^{-1}(x)$ *does* ***not*** *equal* $\frac{1}{f(x)}$.

Informally, we define the inverse of a function as the relation we obtain by switching the domain and range of the function. Because of this definition, you can find an inverse by switching the roles of x and y in an equation. For example, consider the function $g(x) = 2x$. This is the line $y = 2x$. If we switch x and y, we get the equation $x = 2y$. Dividing both sides by 2, we get $y = 1/2\ x$. Therefore the functions $g(x) = 2x$ and $y = 1/2\ x$ are inverses. Using function notation, we can write $y = 1/2\ x$ as $g^{-1}(x) = 1/2\ x$.

Example 1: Find the inverse of each function

Table 4.3:

a. $f(x) = 5x$ - 8	b. $f(x) = x^3$.

Solution:

a. First write the function using " y =" notation, then interchange x and y:

$f(x) = 5x - 8 \rightarrow y = 5x - 8 \rightarrow x = 5y - 8$

Then isolate y:

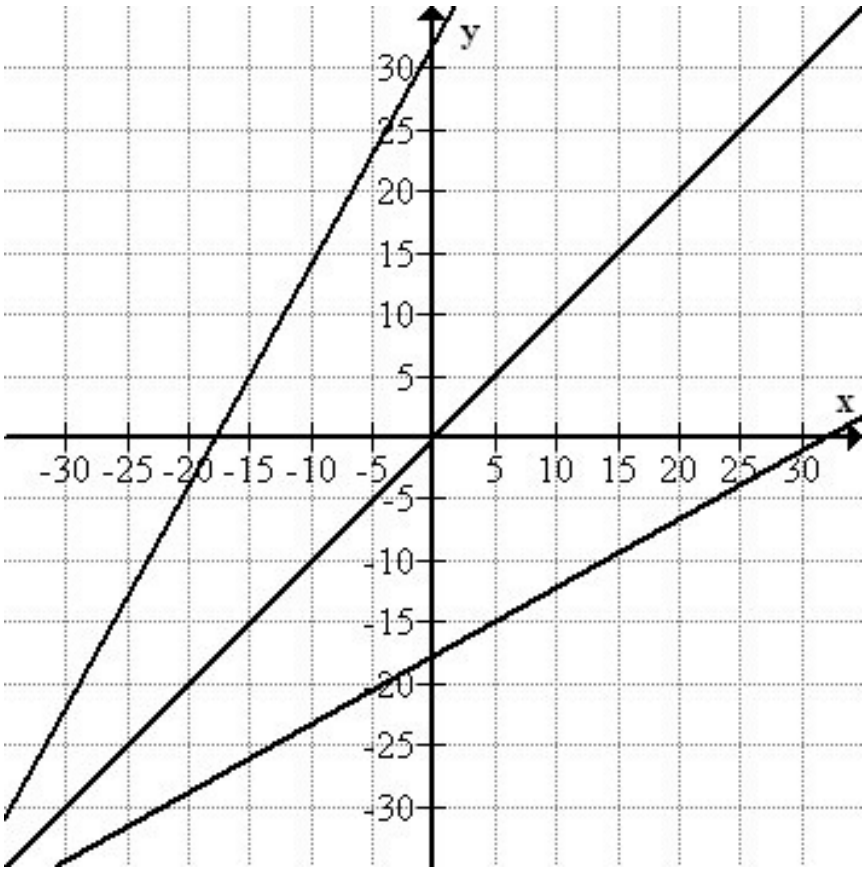
$$x = 5y - 8$$
$$x + 8 = 5y$$
$$y = \frac{1}{5}x + \frac{8}{5}$$

b.

Table 4.4:

First write the function using " y= ":	$f(x) = x^3$
	$y = x^3$
Now interchange x and y	$x = y^3$
Now isolate y:	$y = \sqrt[3]{x}$

Because of the definition of inverse, the graphs of inverses are reflections across the line $y = x$. The graph below shows $t(x) = \frac{5}{9}(x - 32)$ and $f(x) = \frac{9}{5}x + 32$ on the same graph, along with the reflection line $y = x$.

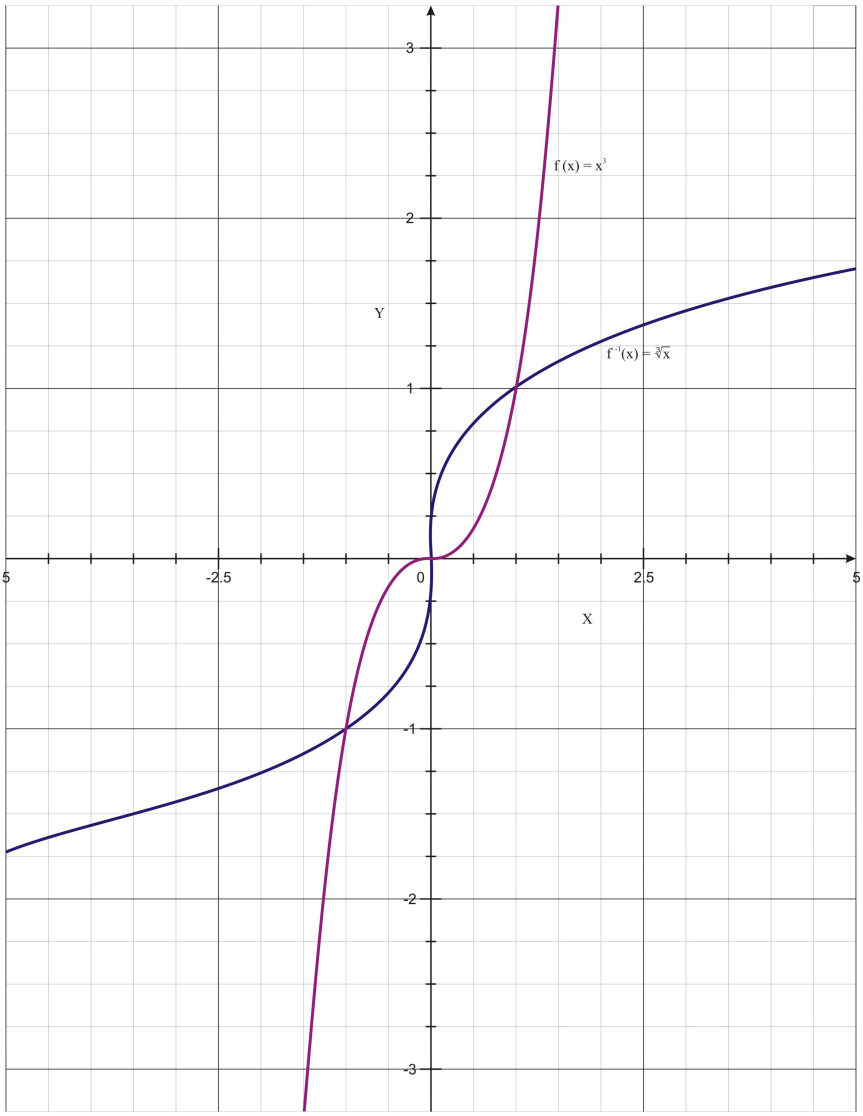


A note about graphing with software or a graphing calculator: if you look at the graph above, you can see that the lines are reflections over the line $y = x$. However, if you do not view the graph in a window that shows equal scales of the x- and y-axes, the graph might not look like this.

Before continuing, there are two other important things to note about inverses. First, remember that the "' -1'" is not an exponent, but a symbol that represents an inverse. Second, not every function has an inverse that is a function. In the examples we have considered so far, we inverted a function, and the resulting relation was also a function. However, some functions are not **invertible**; that is, following the process of "inverting" them does not produce a relation that is a function. We will return to this issue below when we examine domain and range of functions and their inverses. First we will look at a set of functions that *are* invertible.

Inverses of 1-to-1 functions

Consider again example 1 above. We began with the function $f(x) = x^3$, and we found the inverse $f^{-1}(x) = \sqrt[3]{x}$. The graphs of these functions are show below.



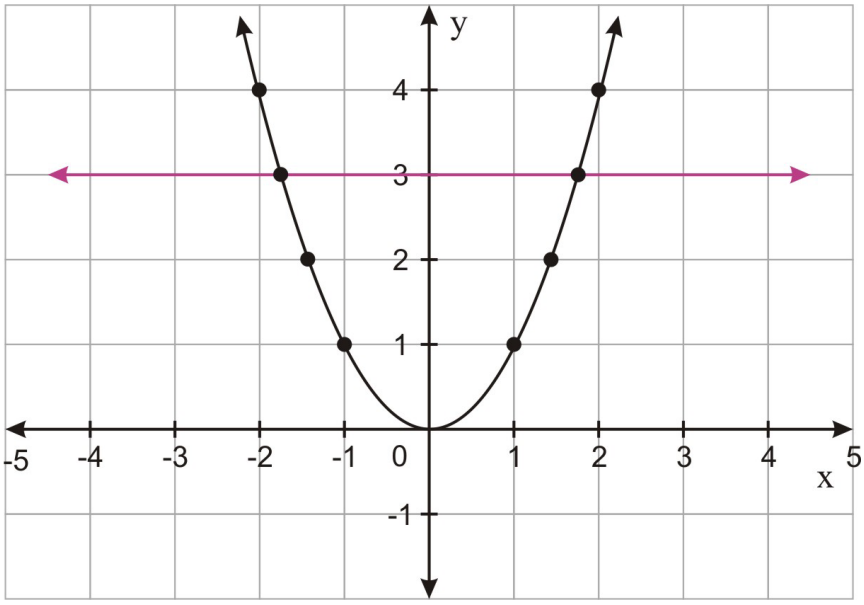
The function $f(x) = x^3$ is an example of a **one-to-one function,** which is defined as follows:

Table 4.5:

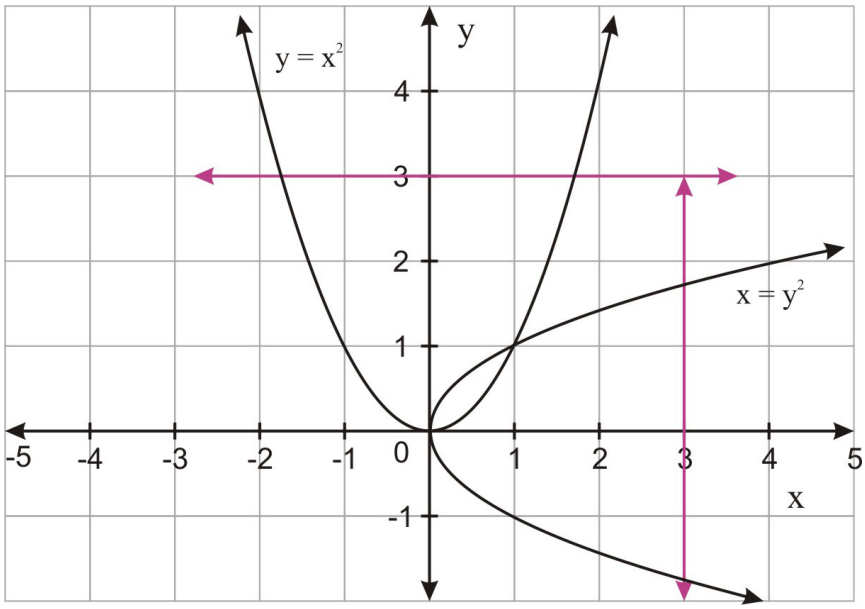
One to one

A function is **one-to-one** if and only if every element of its domain corresponds to *exactly* one element of its range.

The linear functions we examined above are also one-to-one. The function $y = x^2$, however, is not one-to-one. The graph of this function is shown below.



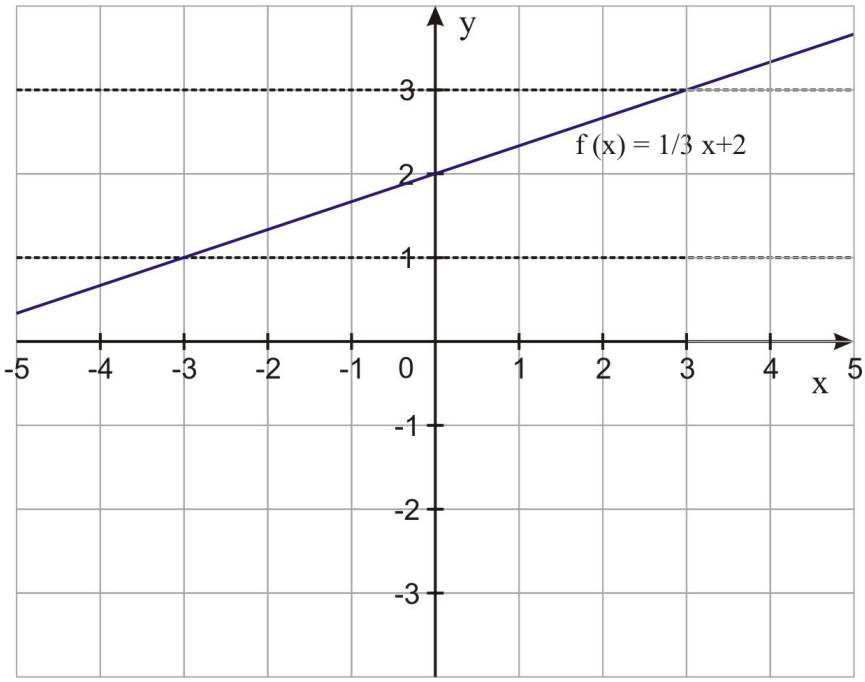
You may recall that you can identify a relation as a function if you draw a vertical line through the graph, and the line touches only one point. Notice then that if we draw a horizontal line through $y = x^2$, the line touches more than one point. Therefore if we inverted the function, the resulting graph would be a reflection over the line $y = x$, and the inverse would not be a function. It fails the vertical line test.



The function $y = x^2$ is therefore *not* a one-to-one function. A function that *is* one-to-one will be invertible. You can determine this graphically by drawing a horizontal line through the graph of the function. For example, if you draw a horizontal line through the graph of $f(x) = x^3$, the line will only touch one point on the graph, no matter where you draw the line.

Example 2: Graph the function $f(x) = \frac{1}{3}x + 2$. Use a horizontal line test to verify that the function is invertible.

Solution: The graph below shows that this function is invertible. We can draw a horizontal line at any y value, and the line will only cross $f(x) = \frac{1}{3}x + 2$ once.



In sum, a one-to-one function is invertible. That is, if we invert a one-to-one function, its inverse is also a function. Now that we have established what it means for a function to be invertible, we will focus on the domain and range of inverse functions.

Domain and range of functions and their inverses

Because of the definition of inverse, a function's domain is its inverse's range, and the inverse's domain is the function's range. This statement may seem confusing without a specific example.

Example 3: State the domain and range of the function and its inverse:

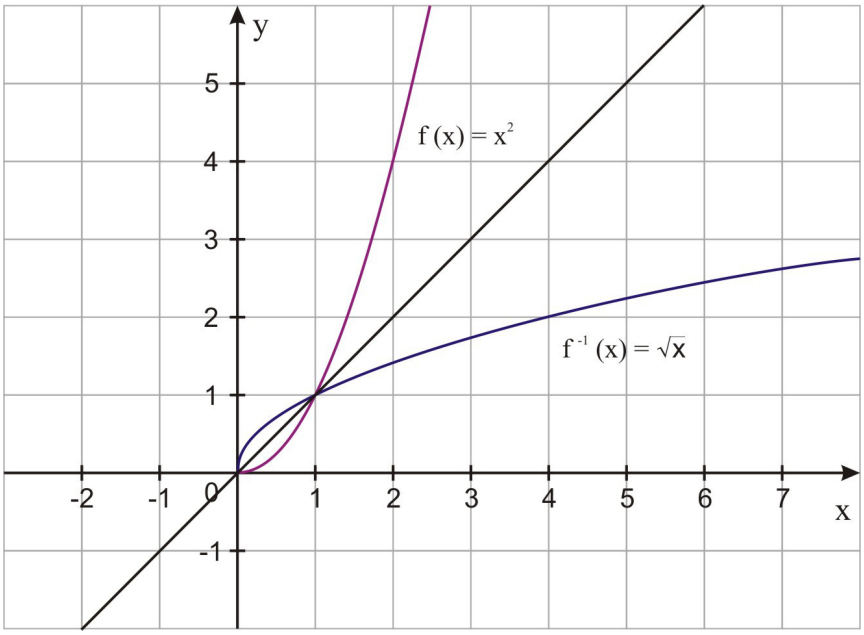
Function: *(1, 2), (2, 5), (3, 7)*

Solution: the inverse of this function is the set of points *(2, 1), (5, 2), (7, 3)*

The domain of the function is *{1, 2, 3}*. This is also the range of the inverse.
The range of the function is *{2, 5, 7}*. This is also the domain of the inverse.

The linear functions we examined previously, as well as $f(x) = x^3$, all had domain and range both equal to the set of all real numbers. Therefore the inverses also had domain and range equal to the set of all real numbers. Because the domain and range were the same for these functions, switching them maintained that relationship.

Also, as we found above, the function $y = x^2$ is not one-to-one, and hence it is not invertible. That is, if we invert it, the resulting relation is not a function. We can change this situation if we define the domain of the function in a more limited way. Let $f(x)$ be a function defined as follows: $f(x) = x^2$, with domain limited to real numbers 0. Then the inverse of the function is the square root function: $f^{-1}(x) = \sqrt{x}$



Example 4: Define the domain for the function $f(x) = (x - 2)^2$ so that f is invertible.

Solution: The graph of this function is a parabola. We need to limit the domain to one side of the parabola. Conventionally in cases like these we choose the positive side; therefore, the domain is limited to real numbers 2.

Inverse functions and composition

In the examples we have considered so far, we have taken a function and found its inverse. We can also analyze two functions and determine whether or not they are inverses. Recall the formal definition from

above:

Two functions $f(x)$ and $g(x)$ are inverses if and only if $f(g(x)) = g(f(x)) = x$.

This definition is perhaps easier to understand if we look at a specific example. Let's use two functions that we have established as inverses: $f(x) = 2x$ and $g(x) = 1/2\ x$. Let's also consider a specific x value. Let $x = 8$. Then we have $f(g(8)) = f(1/2 \times 8) = f(4) = 2(4) = 8$. Similarly we could establish that $g(f(8)) = 8$. Notice that there is nothing special about $x = 8$. For any x value we input into f, the same value will be output by the composed functions:

$$f(g(x)) = f\left(\tfrac{1}{2}x\right) = 2\left(\tfrac{1}{2}x\right) = x$$
$$g(f(x)) = g(2x) = \tfrac{1}{2}(2x) = x$$

Example 5: Use composition of functions to determine if $f(x) = 2x + 3$ and $g(x) = 3x - 2$ are inverses.

Solution: The functions are not inverses.

We only need to check one of the compositions:
$f(g(x)) = f(3x - 2) = 2(3x - 2) + 3 = 6x - 4 + 3 = 6x - 1 \quad x$

Lesson Summary

In this lesson we have defined the concept of inverse, and we have examined functions and their inverses, both algebraically and graphically. We established that functions that are one-to-one are invertible, while other functions are not necessarily invertible. (However, we can redefine the domain of a function such that it is invertible.) In the remainder of the chapter we will examine two families of functions whose members are inverses.

Points to Consider

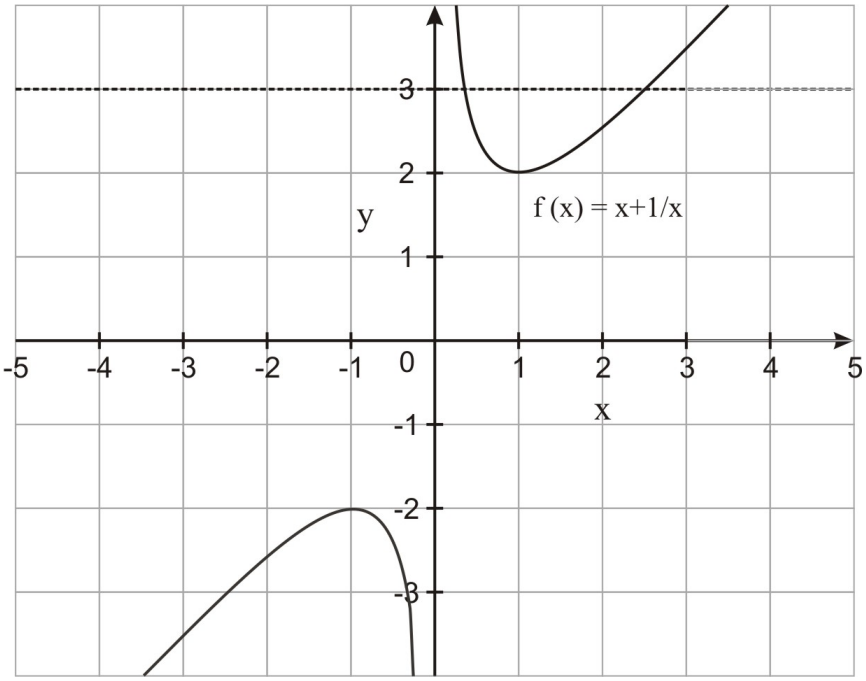
1. Can a function be its own inverse? If so, how?
2. Consider the other function families you learned about in chapter 1. What do their inverses look like?
3. How is the rate of change of a function related to the rate of change of the function's inverse?

Review Questions

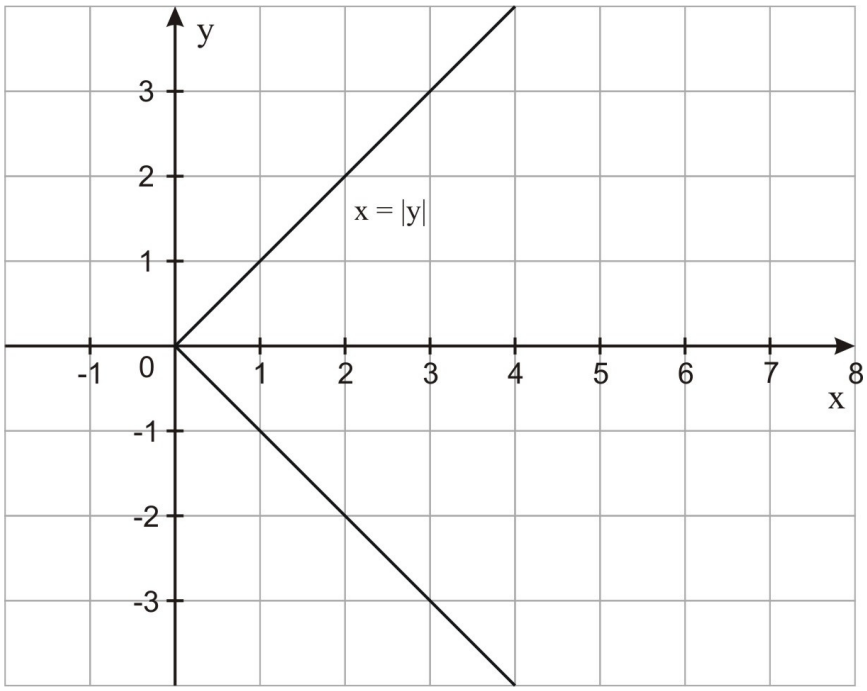
1. Find the inverse of the function $f(x) = \frac{1}{2}x - 7$.
2. Use the horizontal line test to determine if the function $f(x) = x + \frac{1}{x}$ is invertible or not.
3. Use composition of functions to determine if the functions are inverses: $g(x) = 2x - 6$ and $h(x) = \frac{1}{2}x + 3$.
4. Use composition of functions to determine if the functions are inverses: $f(x) = x + 2$ and $p(x) = x - \frac{1}{2}$.
5. Given the function $f(x) = (x + 1)^2$, how should the domain be restricted so that the function is invertible?
6. Consider the function $f(x) = \frac{3}{2}x + 4$. a. Find the inverse of the function.
   b. State the slope of the function and its inverse. What do you notice?
7. Given the function (0, 5), (1, 7), (2, 13), (3, 19) a. Find the inverse of the function.
   b. State the domain and range of the function.
   c. State the domain and range of the inverse.

8. Consider the function $a(x) =$
9. Consider the function $f(x) = c$, where c is a real number. What is the inverse? Is f invertible? Explain.
10. A store sells fabric by the length. Red velvet goes on sale after Valentine's day for \$4.00 per foot. a. Write a function to model the cost of x feet of red velvet.
    b. What is the inverse of this function?
    c. What does the inverse represent?

Review Answers

1. $y = 2x + 14$
2. 

The function is not invertible.
3. The functions are inverses. $g(h(x)) = g\left(\frac{1}{2}x + 3\right) = 2\left(\frac{1}{2}x + 3\right) - 6 = x + 6 - 6 = x$
   $h(g(x)) = h(2x - 6) = \frac{1}{2}(2x - 6) + 3 = x - 3 + 3 = x$
4. The functions are not inverses. $f(p(x)) = (x - \frac{1}{2}) + 2 = x + \frac{3}{2} \neq x$
5. $x \geq -1$
6. a. $y = \frac{2}{3}x - \frac{8}{3}$b. The slope of the function is 3/2 and the slope of the inverse is 2/3. The slopes are reciprocals.
7. a. (5, 0), (7, 1), (13, 2), (19, 3) b. Domain: {0, 1, 2, 3} and Range: {5, 7, 13, 19}
   c. Domain: {5, 7, 13, 19} and Range: {0, 1, 2, 3
8. }
9. a.



b. The function is not invertible. Several ways to justify: the inverse fails the vertical line test; the original function fails the horizontal line test.

10. The function f is a horizontal line with equation $y = c$. The domain is the set of all real numbers, and the range is the single value c. Therefore the inverse would be a function whose domain is c and the range is all real numbers. This is the vertical line $x = c$. This is not a function. So $f(x) = c$ is not invertible.
11. a. $C(x) = 4x$b. $C^{-1}x = \frac{1}{4}x$
    c. The inverse function tells you the number of feet you bought, given the amount of money you spent.

Vocabulary

Inverse The inverse of a function is the relation obtained by interchanging the domain and range of a function.

Invertible A function is invertible if its inverse is a function.

One-to-one A function is one-to-one if every element of its domain is paired with exactly one element of its range.

4.2 Exponential Functions

Learning objectives

- Evaluate exponential expressions
- Identify the domain and range of exponential functions
- Graph exponential functions by hand and using a graphing utility
- Solve basic exponential equations

Introduction

In this lesson you will learn about **exponential functions,** a family of functions we have not studied in chapter 1 or chapter 2. In terms of the form of the equation, exponential functions are different from the other function families because the variable x is in the exponent. For example, the functions $f(x) = 2^x$ and $g(x) = 100(2)5^x$ are exponential functions. This kind of function can be used to model real situations, such as population growth, compound interest, or the decay of radioactive materials. In this lesson we will look at basic examples of these functions, and we will graph and solve exponential equations. This introduction to exponential functions will prepare you to study applications of exponential functions later in this chapter.

Evaluating Exponential Functions

Consider the function $f(x) = 2^x$. When we input a value for x, we find the function value by raising 2 to the exponent of x. For example, if $x = 3$, we have $f(3) = 2^3 = 8$. If we choose larger values of x, we will get larger function values, as the function values will be larger powers of 2. For example, $f(10) = 2^{10} =$ 1,024.

Now let's consider smaller x values. If $x = 0$, we have $f(0) = 2^0 = 1$. If $x =$ -3, we have $f(-3) = 2^{-3} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$. If we choose smaller and smaller x values, the function values will be smaller and smaller fractions. For example, if $x =$ -10, we have $f(-10) = 2^{-10} = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$. Notice that none of the x values we choose will result in a function value of 0. (This is the case because the numerator of the fraction will always be 1.) This tells us that while the domain of this function is the set of all real numbers, the range is limited to the set of positive real numbers. In the following example, you will examine the values of a similar function.

Example 1: For the function $g(x) = 3^x$, find $g(2)$, $g(4)$, $g(0)$, $g(-2)$, $g(-4)$.

Solution:

$g(2) = 3^2 = 9$

$g(4) = 3^4 = 81$

$g(0) = 3^0 = 1$

$g(-2) = 3^{-2} = \frac{1}{3^2} = \frac{1}{9}$

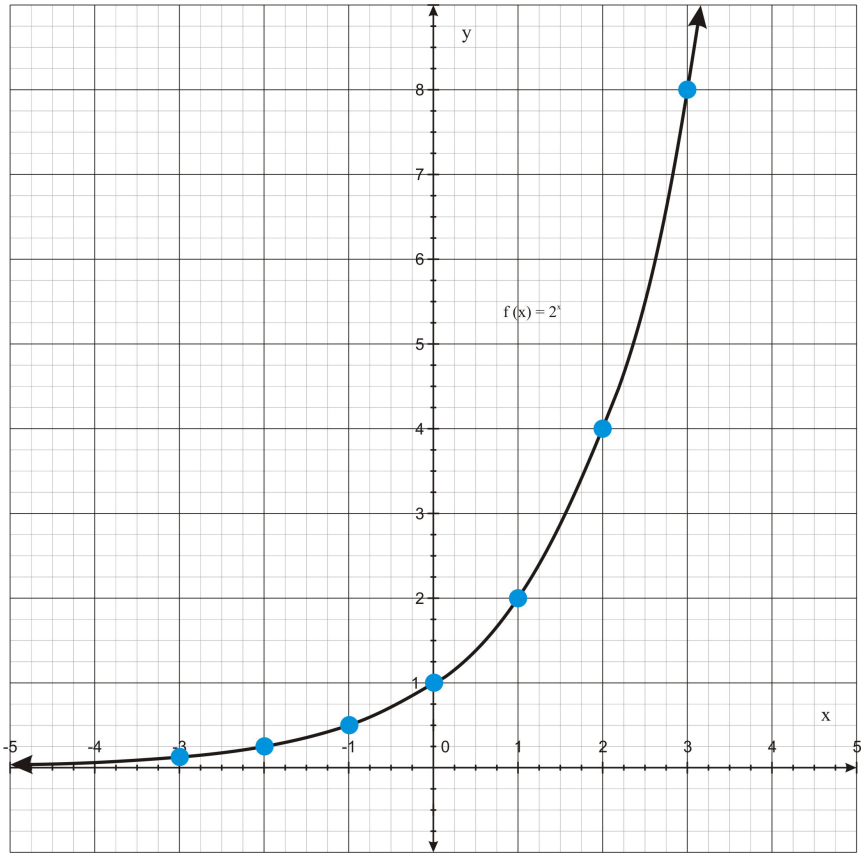
$g(-4) = 3^{-4} = \frac{1}{3^4} = \frac{1}{81}$

The values of the function $g(x) = 3^x$ behave much like those of $f(x) = 2^x$: if we choose larger values, we get larger and larger function values. If $x = 0$, the function value is 1. And, if we choose smaller and smaller x values, the function values will be smaller and smaller fractions. Also, the range of $g(x)$ is limited to positive values.

In general, if we have a function of the form $f(x) = a^x$, where a is a positive real number, the domain of the function is the set of all real numbers, and the range is limited to the set of positive real numbers. This restricted domain will result in a specific shape of the graph.

Graphing basic exponential functions

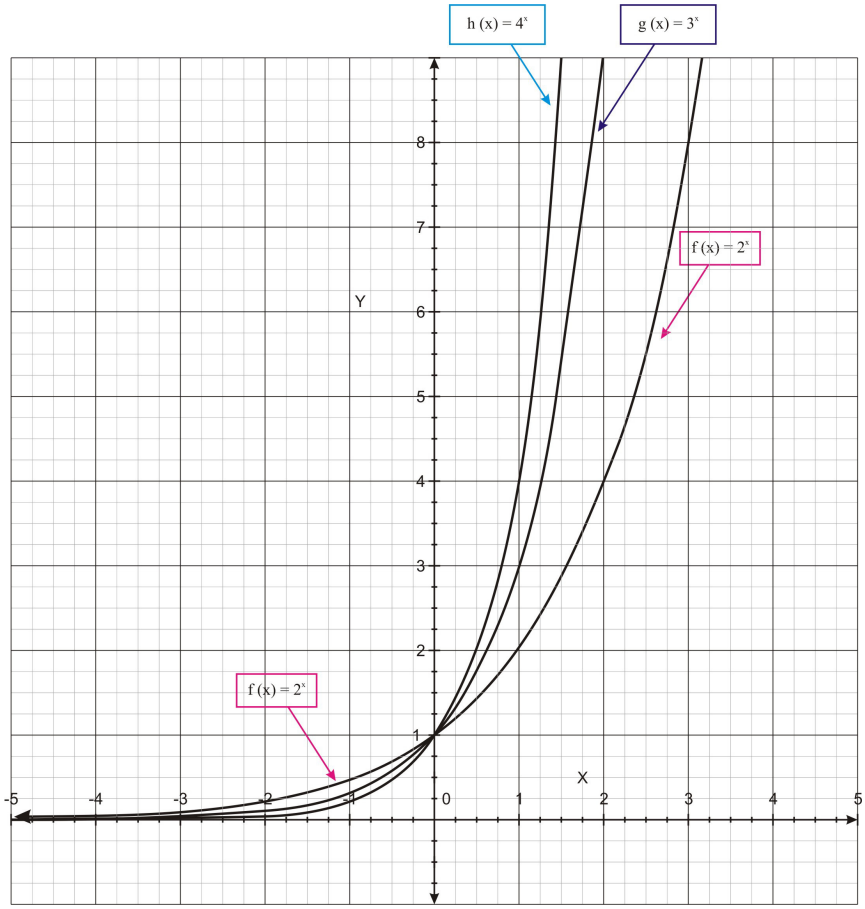
Let's now consider the graph of $f(x) = 2^x$. Above we found several function values, and we began to analyze the function in terms of large and small values of x. The graph below shows this function, with several points marked in blue.



Notice that as x approaches ∞, the function grows without bound. That is, $\lim_{x\to\infty}(2^x) = \infty$. However, if x approaches $-\infty$, the function values get closer and close to 0. That is, $\lim_{x\to-\infty}(2^x) = 0$. Therefore the function is asymptotic to the x-axis. This is the graphical result of the fact that the range of the function is limited to positive y values. Now let's consider the graph of $g(x) = 3^x$ and $h(x) = 4^x$.

Example 2: Use a graphing utility to graph $f(x) = 2^x$, $g(x) = 3^x$ and $h(x) = 4^x$. How are the graphs the same, and how are they different?

Solution: $f(x) = 2^x$, $g(x) = 3^x$ and $h(x) = 4^x$ are shown together below.



The graphs of the three functions have the same overall shape: they have the same end behavior, and they all contain the point (0, 1). The difference lies in their rate of growth. Notice that for positive x values, $h(x) = 4^x$ grows the fastest and $f(x) = 2^x$ grows the slowest. The function values for $h(x) = 4^x$ are highest and the function values for $f(x) = 2^x$ are the lowest for any given value of x. For negative x values, the relationship changes: $f(x) = 2^x$ has the highest function values of the three functions.

Now that we have examined these three parent graphs, we will graph using shifts, reflections, stretches and compressions.

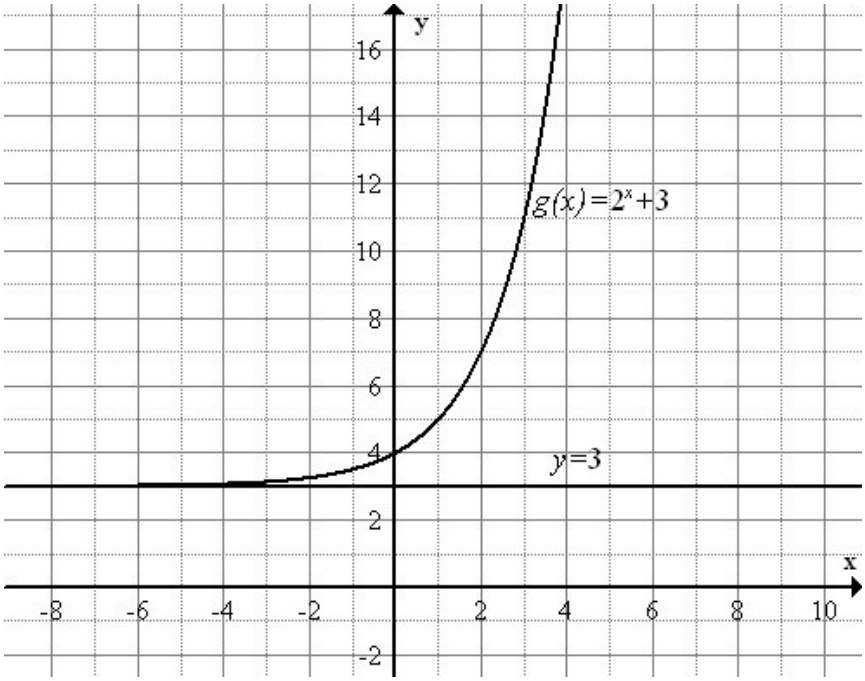
Graphing exponential functions using transformations.

Above we graphed the function $f(x) = 2^x$. Now let's consider a related function: $g(x) = 2^x + 3$. Every function value will be a power of 2, plus 3. The table below shows several values for the function:

Table 4.6:

x	$g(x) = 2^x + 3$
-2	$2^{-2} + 3 = \frac{1}{4} + 3 = 3\frac{1}{4}$
-1	$2^{-1} + 3 = \frac{1}{2} + 3 = 3\frac{1}{2}$
0	$2^0 + 3 = 1 + 3 = 4$
1	$2^1 + 3 = 2 + 3 = 5$
2	$2^2 + 3 = 4 + 3 = 7$
3	$2^3 + 3 = 8 + 3 = 11$

The function values follow the same kind of pattern as the values for $f(x) = 2^x$. However, because every function value is 3 more than a power of 2, the horizontal asymptote of the function is the line $y = 3$. The graph of this function and the horizontal asymptote are shown below.



From your study of transformation of functions in chapter 1, you should recognize the graph of $g(x) = 2^x + 3$ as a vertical shift of the graph of $f(x) = 2^x$. In general, we can produce a graph of an exponential function with base 2 if we analyze the equation of the function in terms of transformations. The table below summarizes the different kinds of transformations of $f(x) = 2^x$. The issue of stretching will be discussed further below the table.

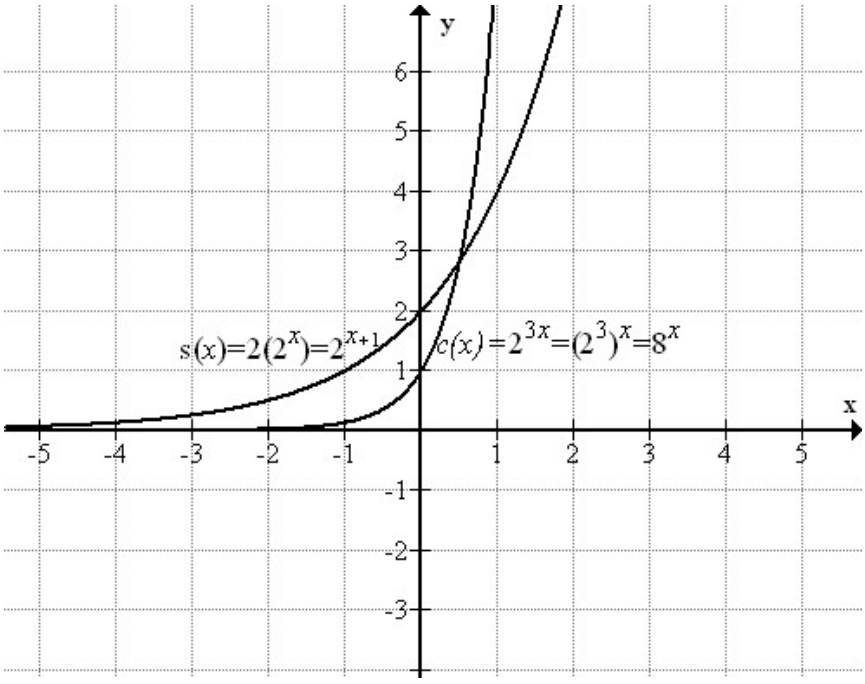
Table 4.7:

Equation	Relationship to $f(x)=2^x$	Range
$g(x) = \frac{2^x}{2^a} = 2^{x-a}$, for $a > 0$	Obtain a graph of g by shifting the graph of f a units to the right.	$y > 0$
$g(x) = 2^a \cdot 2^x = 2^{a+x}$, for $a > 0$	Obtain a graph of g by shifting the graph of f a units to the left.	$y > 0$
$g(x) = 2^x + a$, for $a > 0$	Obtain a graph of g by shifting the graph of f up a units.	$y >$ a
$g(x) = 2^x - a$, for $a > 0$	Obtain a graph of g by shifting the graph of f down a units.	$y >$ a
$g(x) = a(2^x)$, for $a > 0$	Obtain a graph of g by vertically stretching the graph of f *by a factor of a.*	$y > 0$
$g(x) = 2^{ax}$, for $a > 0$	Obtain a graph of g by horizontally compressing the graph of f *by a factor of a.*	$y > 0$
$g(x) = -2^x$	Obtain a graph of g by reflecting the graph of f over the x-axis.	$y > 0$
$g(x) = 2^{-x}$	Obtain a graph of g by reflecting the graph of f over the y-axis.	$y > 0$

Table 4.7: (continued)

Equation	Relationship to $f(x)=2^x$	Range

As was discussed in chapter 1, a stretched graph can also be seen as a compressed graph. This is not the case for exponential functions because of the x in the exponent. Consider the function $s(x) = 2(2^x)$ and $c(x) = 2^{3x}$. The first function represents a vertical stretch of $f(x) = 2^x$ by a factor of 2. The second function represents a compression of $f(x) = 2^x$ by a factor of 3. The function c(x) is actually the same as another parent function: $c(x) = 2^{3x} = (2^3)^x = 8^x$. The function s(x) is actually the same as a shift of $f(x) = 2^x$: $s(x) = 2(2^x) = 2^1 \times 2^x = 2^{x+1}$. The graphs of s and c are shown below. Notice that the graph of c has a y-intercept of 1, while the graph of s has a y-intercept of 2:

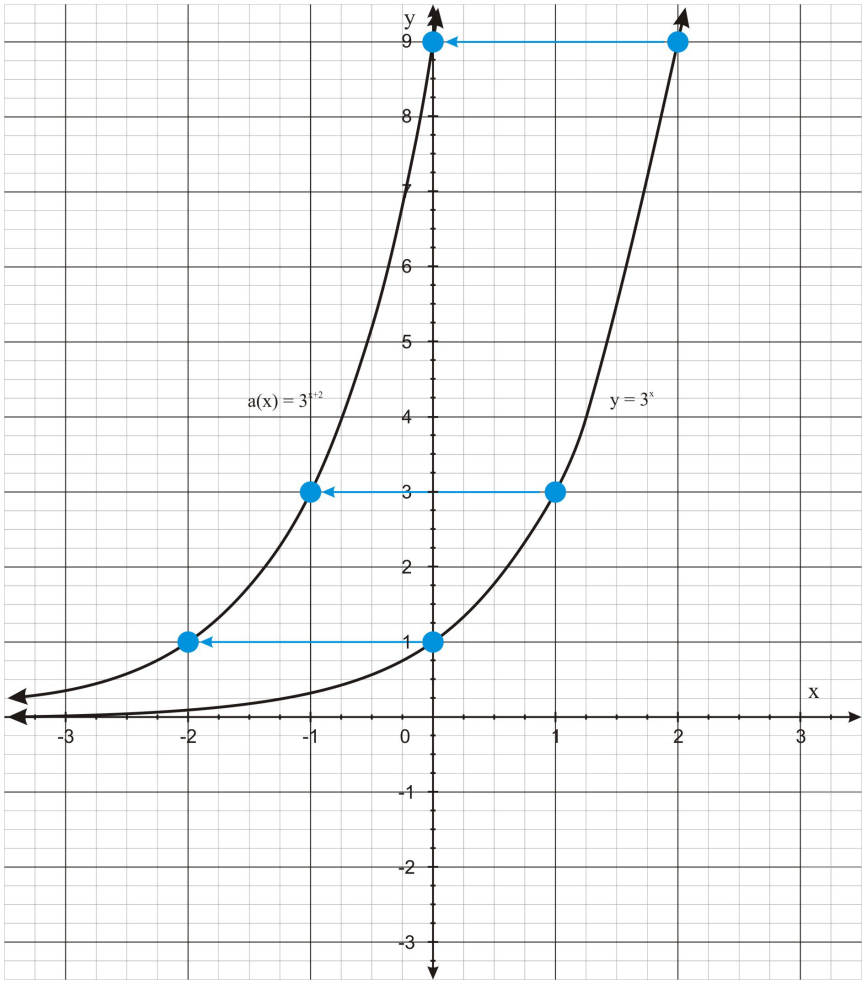


Example 3: Use transformations to graph the functions (a) $a(x) = 3^{x+2}$ and (b) $b(x) = -3^x + 4$

Solution:

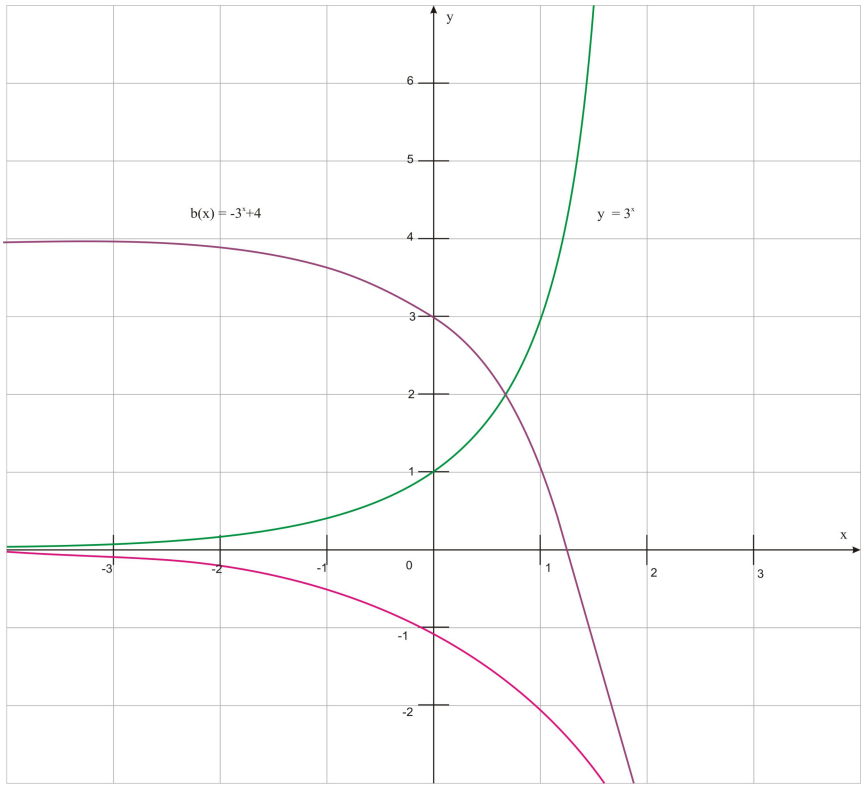
a. $a(x) = 3^{x+2}$

This graph represents a shift of $y = 3^x$ two units to the left. The graph below shows this relationship between the graphs of these two functions:



b. $b(x) = -3^x + 4$

This graph represents a reflection over the y-axis and a vertical shift of 4 units. You can produce a graph of $b(x)$ using three steps: sketch $y = 3^x$, reflect the graph over the x-axis, and then shift the graph up 4 units. The graph below shows this process:



While you can always quickly create a graph using a graphing utility, using transformations will allow you to sketch a graph relatively quickly on your own. If we start with a parent function such as $y=3^x$, you can quickly plot several points: (0, 1), (2, 9), (-1, 1/3), etc. Then you can transform the graph, as we did in the previous example.

Notice that when we sketch a graph, we choose x values, and then use the equation to find y values. But what if we wanted to find an x value, given a y value? This requires solving exponential equations.

Solving exponential equations

Solving an exponential equation means determining the value of x for a given function value. For example, if we have the equation $2^x = 8$, the solution to the equation is the value of x that makes the equation a true statement. Here, the solution is $x = 3$, as $2^3 = 8$.

Consider a slightly more complicated equation $3\ (2^{x\ +\ 1}) = 24$. We can solve this equation by writing both sides of the equation as a power of 2:

$$
\begin{aligned}
&3(2^{x+1}) = 24 \\
&\tfrac{3(2^{x+1})}{3} = \tfrac{24}{3} \\
&2^{x+1} = 8 \\
&2^{x+1} = 2^3
\end{aligned}
$$

To solve the equation now, recall a property of exponents: if $b^x = b^y$, then $x = y$. That is, if two powers of the same base are equal, the exponents must be equal. This property tells us how to solve:

$$
\begin{aligned}
&2^{x+1} = 2^3 \\
&\Rightarrow x + 1 = 3
\end{aligned}
$$

$x = 2$

Example 4: Solve the equation $5^{6x + 10} = 25^{x - 1}$

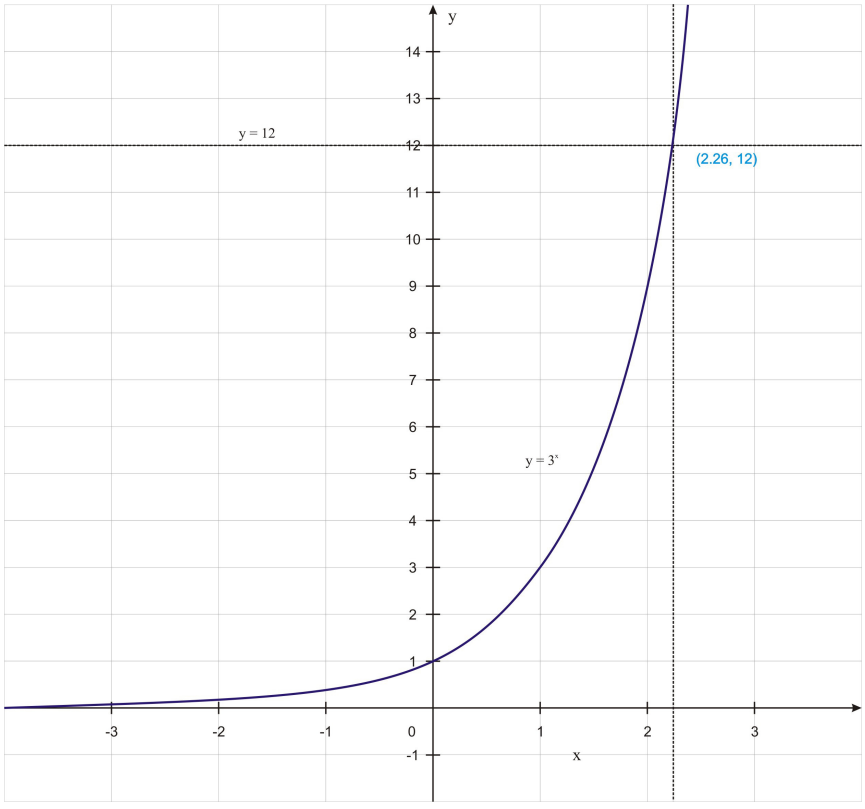
Solution: Use the same technique as shown above:

$$\begin{aligned} 5^{6x + 10} &= 25^{x - 1} \\ 5^{6x + 10} &= (5^2)^{x - 1} \\ 5^{6x + 10} &= 5^{2x - 2} \\ \Rightarrow 6x + 10 &= 2x - 2 \\ 4x + 10 &= -2 \\ 4x &= -12 \\ x &= -3 \end{aligned}$$

In both of the examples of solving equations, it was possible to solve because we could write both sides of the equations as a power of the same exponent. But what if that is not possible?

Consider for example the equation $3^x = 12$. If you try to figure out the value of x by considering powers of 3, you will quickly discover that the solution is not a whole number. Later in the chapter we will study techniques for solving more complicated exponential equations. Here we will solve such equations using graphs.

Consider the function $y = 3^x$. We can find the solution to the equation $3^x = 12$ by finding the intersection of $y = 3^x$ and the horizontal line $y = 12$. Using a graphing calculator's intersection capability, you should find that the solution is approximately $x = 2.26$.



Example 5: Use a graphing utility to solve each equation:

Table 4.8:

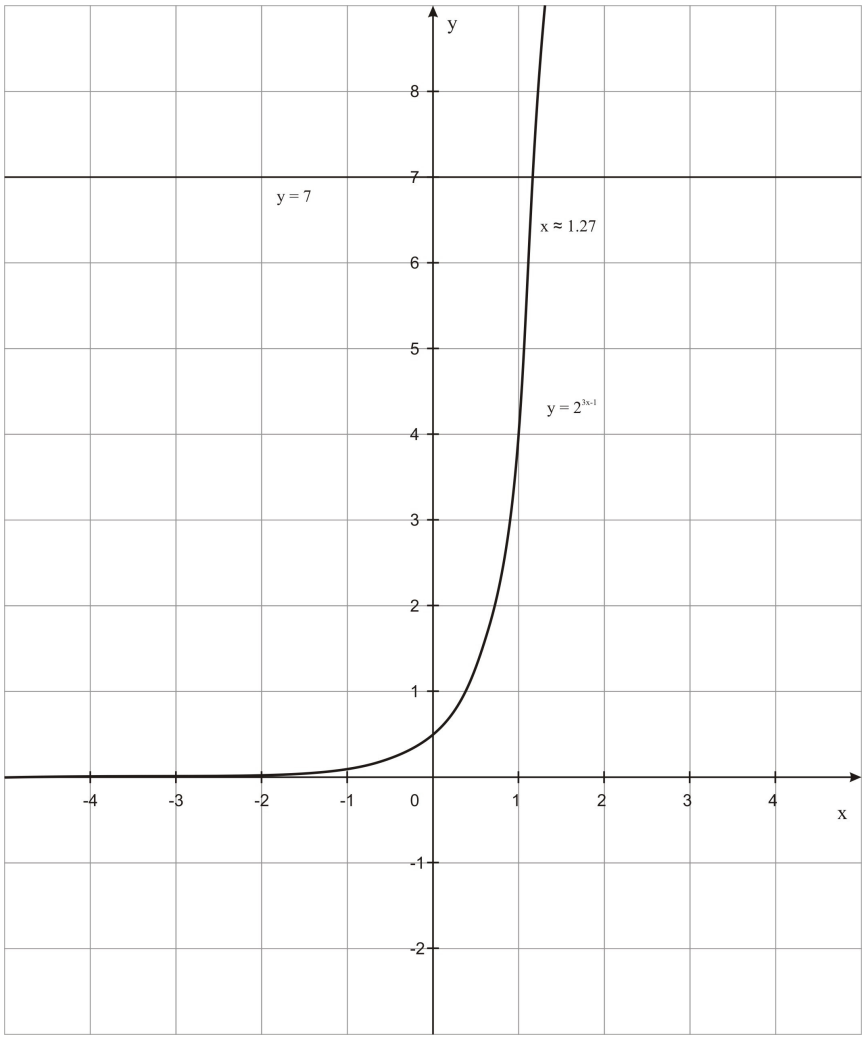
a. $2^{3x - 1} = 7$	b. $6^{-4x} = 2^{8x - 5}$

Solution:

a. $2^{3x - 1} = 7$

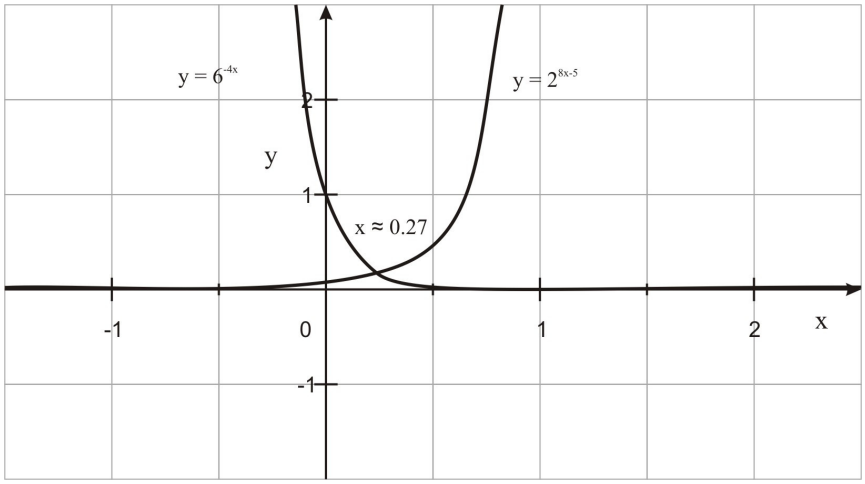
Graph the function $y = 2^{3x - 1}$ and find the point where the graph intersects the horizontal line $y = 7$.

The solution is $x \approx 1.27$



b. $6^{-4x} = 2^{8x - 5}$

Graph the functions $y = 6^{-4x}$ and $y = 2^{8x - 5}$ and find their intersection point.



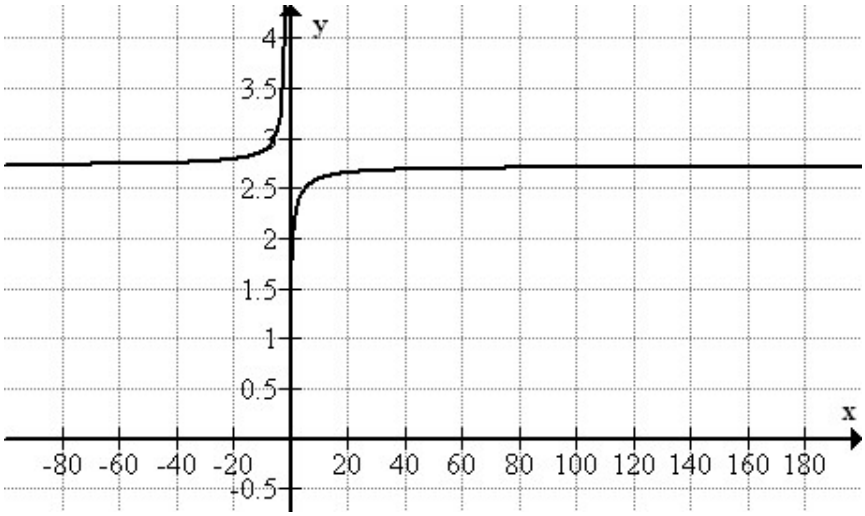
The solution is approximately $x \quad 0.27$. (Your graphing calculator should show 9 digits: 0.272630365.)

In the examples we have considered so far, the bases of the functions have been positive integers. Now we will examine a sub-family of exponential functions with a special base: the number e.

The number e and the function $y = e^x$

In your previous studies of math, you have likely encountered the number . The number e is much like . First, both are **irrational** numbers: they cannot be expressed as fractions. Second, both numbers are **transcendental:** they are not the solution of any polynomial with rational coefficients.

Like , mathematicians found e to be a natural constant in the world. One way to "discover" e is to consider the function $f(x) = \left(1 + \frac{1}{x}\right)^x$. The graph of this function is shown below.



Notice that as x approaches infinity, the graph of the function approaches a horizontal asymptote around $y = 2.7$. If you examine several function values, you will see that the limit is not exactly 2.7:

Table 4.9:

x	y
0	(not defined)
1	2
2	2.25
5	2.48832

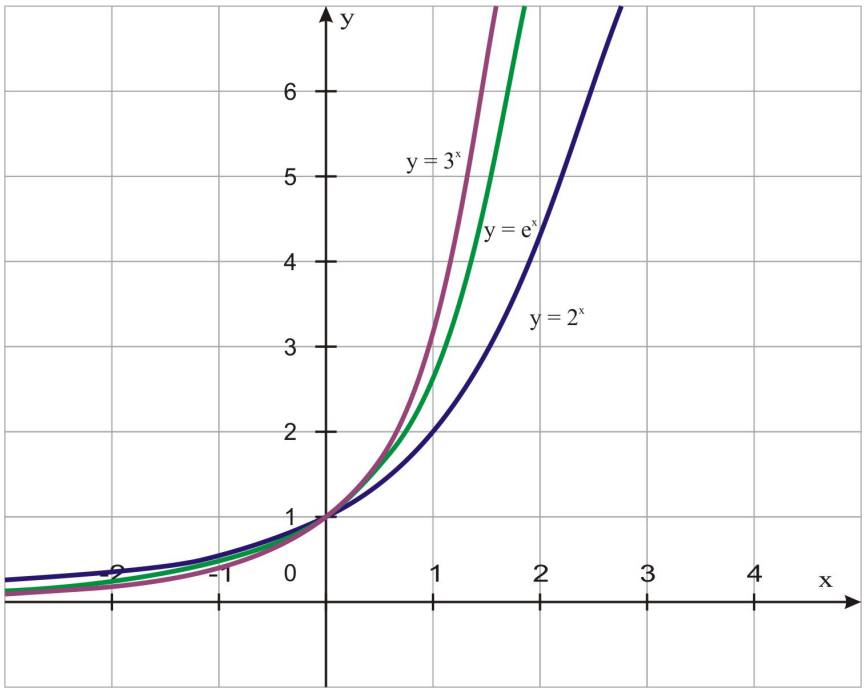
Table 4.9: (continued)

x	y
10	2.5937424601
50	2.69158802907
100	2.70481382942
1000	2.71692393224
5000	2.7180100501
10,000	2.71814592683
50,000	2.7181825464614

Around $x = 100$, the function values pass 2.7, but they will never reach 2.8. The limit of the function as x approaches infinity is the constant e. The value of e is approximately 2.71828182845904523536. Again, like , we have to approximate the value of e because it is irrational.

The number e is used as the base of functions that can be used to model situations that involve growth or decay. For example, as you will learn later in the chapter, one method of calculating interest on a bank account or investment uses this number. Here we will examine the function $y = e^x$ in order to verify that its graph is similar to the other exponential functions we have graphed.

The graph below shows $y = e^x$, along with $y = 2^x$ and $y = 3^x$.



The graph of $y = e^x$ (in green) has the same shape as the graphs of the other exponential functions. It sits in between the graphs of the other two functions, and notice that the graph is closer to $y = 3^x$ than to $y = 2^x$. All three graphs have the same y-intercept: (0, 1). Thus the graph of this function is clearly a member of the same family, even though the base of the function is an irrational number.

Lesson Summary

This lesson has introduced the family of exponential functions. We have examined values of functions, towards understanding the behavior of graphs. In general, exponential functions have a horizontal asymptote, though one end of the function increases (or decreases, if it is a reflection) without bound.

In this lesson we have graphed these functions, solved certain exponential equations using our knowledge of exponents, and solved more complicated equations using graphing utilities. We have also examined the function $y = e^x$, which is a special member of the exponential family. In the coming lessons you will continue to learn about exponential functions, including the inverses of these functions, applications of these functions, and solving more complicated exponential equations using algebraic techniques.

Points to Consider

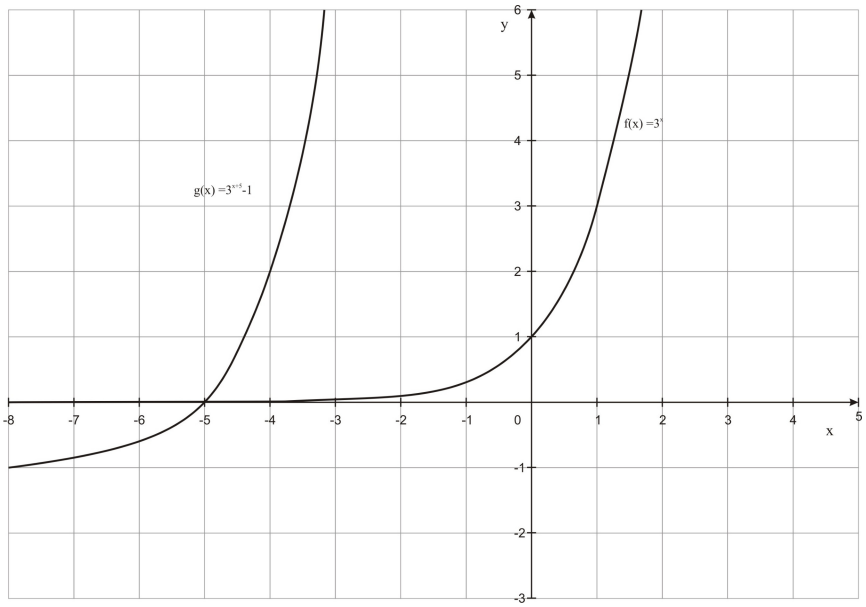
1. Why do exponential functions have horizontal asymptotes and not vertical asymptotes?
2. What would the graph of the inverse of an exponential function look like? What would its domain and range be?
3. How could you solve or approximate a solution to an exponential equation without using a graphing calculator?

Review Questions

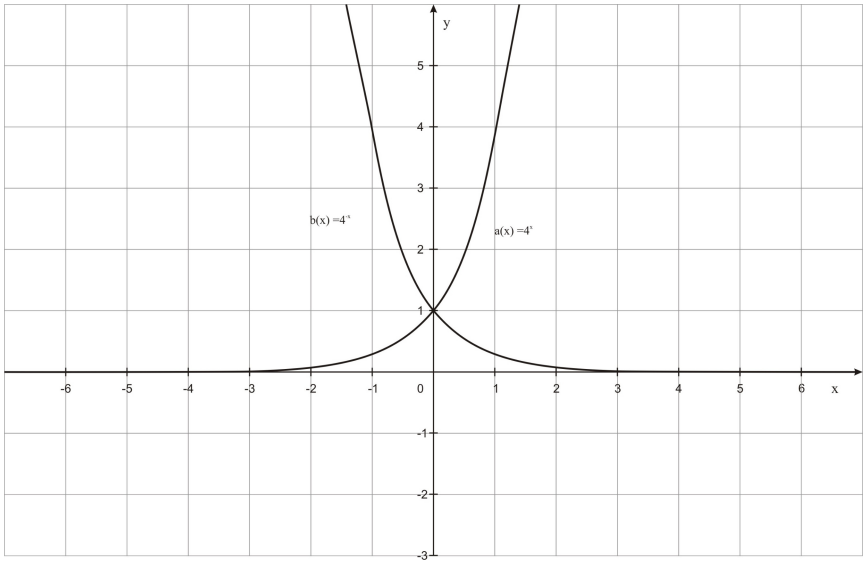
1. For the function $f(x) = 2^{3x - 1}$, find $f(0)$, $f(2)$, and $f(-2)$.
2. Graph the functions $f(x) = 3^x$ and $g(x) = 3^{x + 5} - 1$. State the domain and range of each function.
3. Graph the functions $a(x) = 4^x$ and $b(x) = 4^{-x}$. State the domain and range of each function.
4. Graph the function $h(x) = -2^{x - 1}$ using transformations. How is h related to $y = 2^x$?
5. Solve the equation: $5^{2x + 1} = 25^{3x}$
6. Solve the equation: $4^{x2} + 1 = 16^x$
7. Use a graph to find an approximate solution to the equation $3^x = 14$
8. Use a graph to find an approximate solution to the equation $2^{-x} = 7^{2x + 9}$
9. Sketch a graph of the function $f(x) = 3^x$ and its inverse. (Hint: You can graph the inverse by reflecting a function across the line $y=x$.) Is f one-to-one?
10. Consider the following situation: you inherited a collection of 125 stamps from a relative. You decided to continue to build the collection, and you vowed to double the size of the collection every year. a. Write an exponential function to model the situation. (The input of the function is the number of years since you began building the collection, and the output is the size of the collection.)
    b. Use your model to determine how long it will take to have a collection of 10,000 stamps.

Review Answers

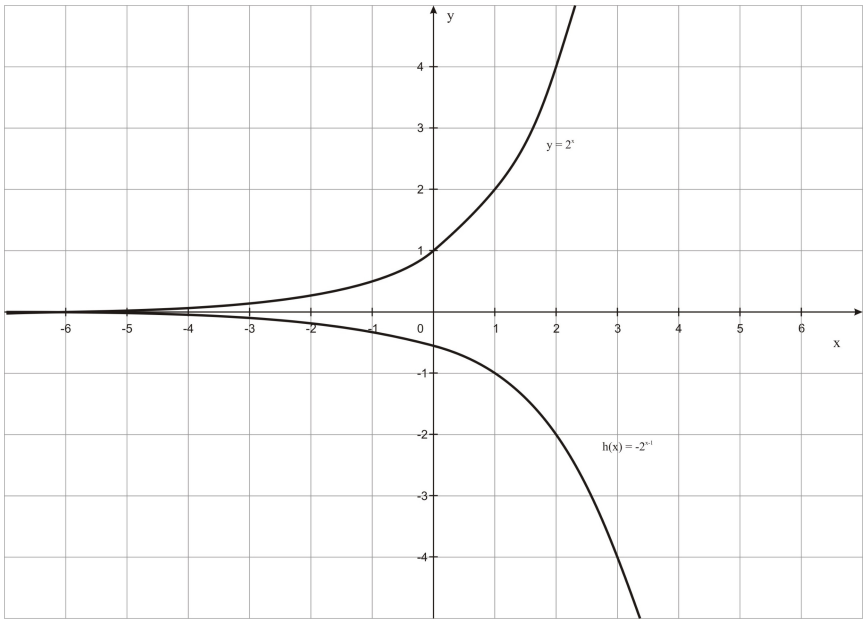
1. f(0)=1/2 , f(2)= 32, f(-2)= 1/128
2. The domain if both functions is the set of all real numbers. The range of f is the set of all real numbers > 0.
   The range of g is the set of all real numbers > -1



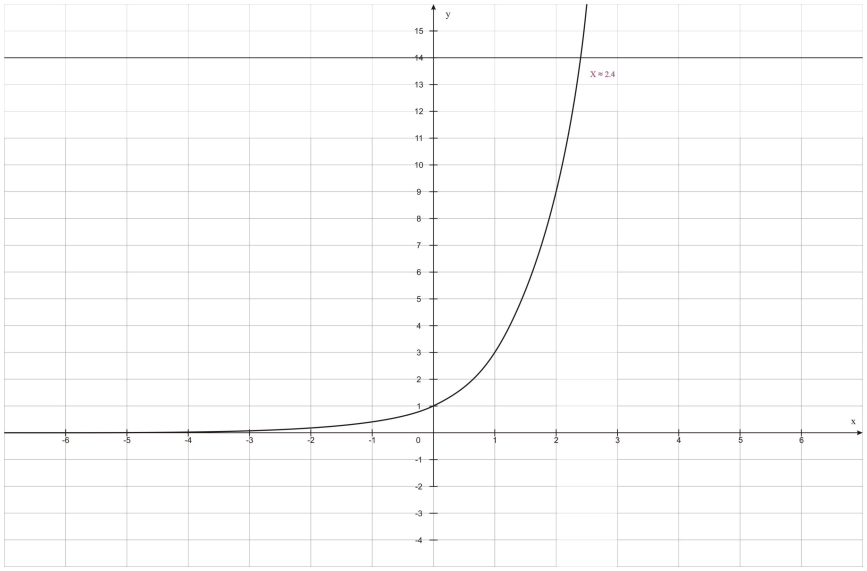
3. The domain of both functions is the set of all real numbers. The range of both functions is the set of all real numbers $\quad 0$.

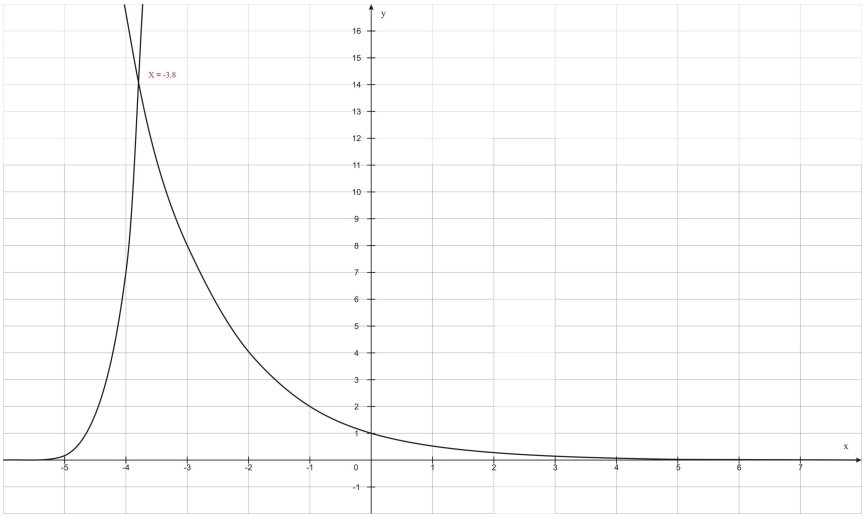


4. The function h represents a reflection over the y axis, and a horizontal shift 1 unit to the right.



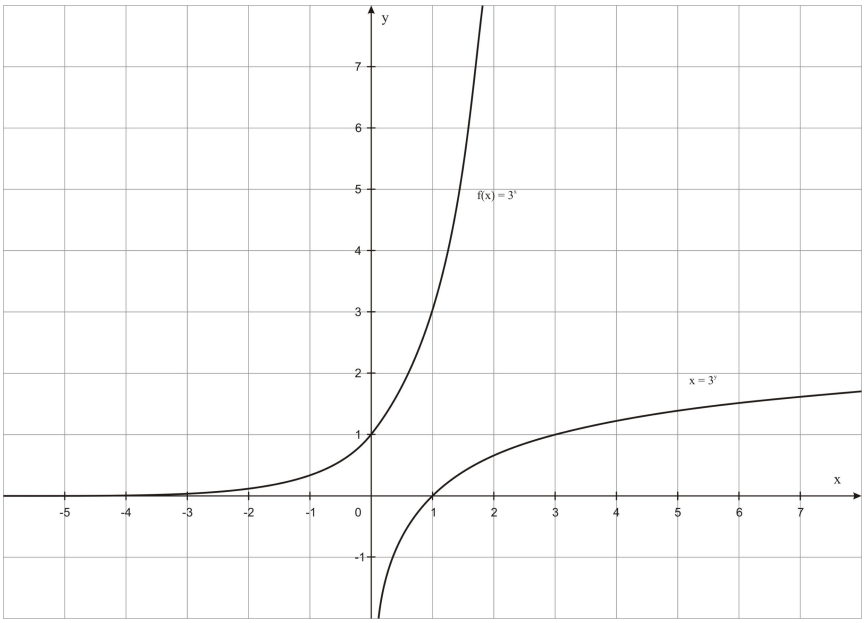
5. $5^{2x+1} = 25^{3x}5^{2x+1} = 5^{6x}$
   $\Rightarrow 2x+ 1 = 6x$
   $4x = 1$
   $x = 1/4$
6. $4^{x2} + 1 = 16^{x}4^{x2} + 1 = 4^{2x}$
   $\Rightarrow x^2 + 1 = 2x$
   x^2 - $2x + 1 = 0$
   $(x$ - $1)$ $(x$ - $1) = 0$
   $x = 1$
7. x 2.4



8. 

x -3.8

9. f is a one-to-one function.



10. a. $S(t) = 125\ (2^{t})$b. About 6.35 years

Vocabulary

e The number e is a transcendental number, often referred to as Euler's constant. Several mathematicians are credited with early work on e. Euler was the first to use this letter to represent the constant. The value of e is approximately 2.71828. The exact value is $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{x}$.

Exponential Function An exponential function is a function for which the input variable x is in the exponent of some base b, where b is a real number.

Irrational number An irrational number is a number that cannot be expressed as a fraction of two integers.

Transcendental number A transcendental number is a number that is not a solution to any non-zero polynomial with rational roots.

The number is a transcendental number. It is the ratio of the circumference to the diameter in any circle.

4.3 Logarithmic Functions

Learning objectives

- Translate numerical and algebraic expressions between exponential and logarithmic form.
- Evaluate logarithmic functions.
- Determine the domain of logarithmic functions.
- Graph logarithmic functions.
- Solve logarithmic equations.

Introduction

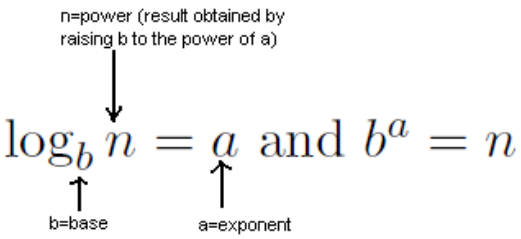
In the previous lesson we examined exponential expressions and functions. Now we will consider another representation for the same relationships involved in exponential expressions and functions.

Consider the function $y = 2^x$. Every x-value of this function is an **exponent**. Every y-value is a power of 2. As you learned in lesson 1, functions that are one-to-one have inverses that are functions. This is the case with exponential functions. If we take the inverse of $y = 2^x$ (by interchanging the domain and range) we obtain this equation: $x = 2^y$. In order to write this equation such that y is expressed as a function of x, we need a different notation.

The solution to this problem is found in the **logarithm**. John Napier originally introduced the logarithm to 17$^{\text{th}}$ century mathematicians as a technique for simplifying complicated calculations. While today's technology allows us to do most any calculations we could imagine, logarithmic functions continue to be a focus of study in mathematics, as a useful way to work with exponential expressions and functions.

Changing Between Exponential and Logarithmic Expressions

Every exponential expression can be written in logarithmic form. For example, the equation $x=2^y$ is written as follows: $y=log_2 x$. In general, the equation $log_b n = a$ is equivalent to the equation $b^a=n$. That is, b is the **base**, a is the **exponent**, and n is the **power,** or the result you obtain by raising b to the power of a. Notice that the exponential form of an expression emphasizes the power, while the logarithmic form emphasizes the exponent. More simply put, a logarithm (or "log" for short) *is* an exponent.



We can write any exponential expression in logarithmic form.

Example 1: Rewrite each exponential expression as a log expression.

Table 4.10:

a. $3^4=81$	b. $b^{4x}=52$

Solution:

a. In order to rewrite an expression, you must identify its base, its exponent, and its power. The 3 is the base, so it is placed as the subscript in the log expression. The 81 is the power, and so it is placed after the "log". Thus we have: $3^4=81$ is the same as $log_3 81=4$.
To read this expression, we say "the logarithm base 3 of 81 equals 4." This is equivalent to saying " 3 to the 4^{th} power equals 81."
b. The b is the base, and the expression 4x is the exponent, so we have:$log_b 52=4x$. We say, "log base b of 52, equals $4x$."

We can also express a logarithmic statement in exponential form.

Example 2: Rewrite the logarithmic expressions in exponential form.

Table 4.11:

a. $\log_{10}100=2$	b. $\log_b w=5$

Solution:

a. The base is 10, and the exponent is 2, so we have: $10^2=100$
b. The base is b, and the exponent is 5, so we have: $b^5=w$.

Perhaps the most common example of a logarithm is the Richter scale, which measures the magnitude of an earthquake. The magnitude is actually the logarithm base 10 of the amplitude of the quake. That is, $m=\log_{10}A$. This means that, for example, an earthquake of magnitude 4 is 10 *times* as strong as an earthquake with magnitude 3. We can see why this is true of we look at the logarithmic and exponential forms of the expressions: An earthquake of magnitude 3 means $3=\log_{10}A$. The exponential form of this expression is $10^3=A$. Thus the amplitude of the quake is 1,000. Similarly, a quake with magnitude 4 has amplitude 10^4=10,000. We will return to this example in lesson 3.8.

Evaluating Logarithmic Functions

As noted above, a logarithmic function is the inverse of an exponential function. Consider again the function $y=2_x$ and its inverse $x=2^y$. Above, we rewrote the inverse as $y=\log_2 x$. If we want to emphasize the fact that the log equation represents a **function**, we can write the equation as $f(x)=\log_2 x$. To **evaluate** this function, we choose values of x and then determine the corresponding y values, or function values.

Example 3: Evaluate the function $f(x)=\log_2 x$ for the values:

Table 4.12:

a.x=2	b.x=1	c.x=-2

Solution:

a. If x=2 , we have:

Table 4.13:

	$f(x)=\log_2 x$
	$f(2)=\log_2 2$

To determine the value of $\log_2 2$, you can ask yourself: "2 to what power equals 2?" Answering this question is often easy if you consider the exponential form: $2^{?}=2$

The missing exponent is 1. So we have $f(2)$-$\log_2 2=1$

b. If x=1 , we have:

Table 4.14:

	$f(x)=\log_2 x$
	$f(1)=\log_2 1$

As we did in (a), we can consider the exponential form: $2^{?}=1$. The missing exponent is 0. So we have $f(1)=\log^2 1=0$.

c. If x=-2, we have:

Table 4.15:

	$f(x)$	$=\log_2 x$
	f(-2)	$=\log_2$-2

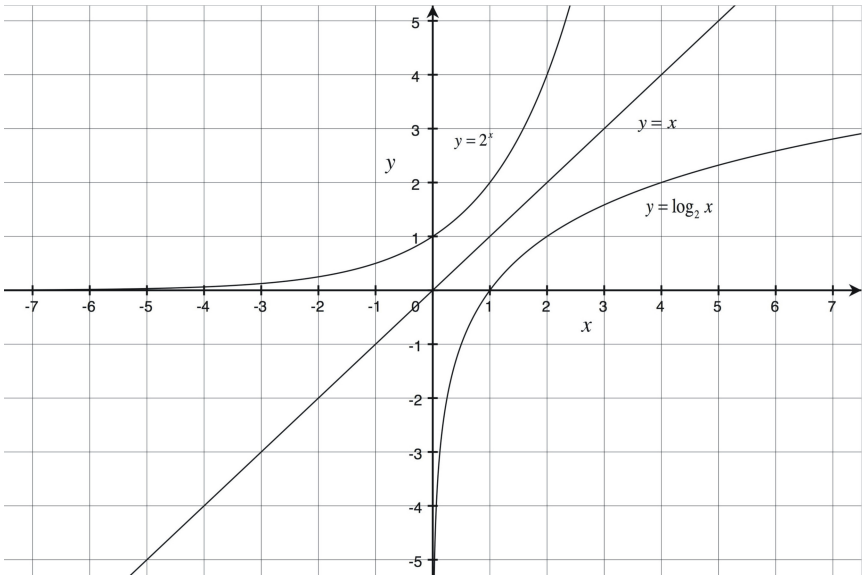
Again, consider the exponential form: $2^{?}$=-2. There is no such exponent. Therefore f(-2)=$\log_2$-2 does not exist.

Example 3c illustrates an important point: there are restrictions on the **domain** of a logarithmic function. For the function $f(x)=\log_2 x$, x cannot be a negative number. Therefore we can state the domain of this function as: "the set of all real numbers greater than 0." Formally, we can write it as a set: $\{x \in | x > 0\}$. In general, the domain of a logarithmic function is restricted to those values that will make the argument of the logarithm non-negative.

For example, consider the function $f(x)=\log_3(x\text{-}4)$. If you attempt to evaluate the function for x values of 4 or less, you will find that the function values do not exist. Therefore the domain of the function is $\{x \in | x > 4\}$. The domain of a logarithmic function is one of several key issues to consider when graphing.

Graphing Logarithmic Functions

Because the function $f(x)=\log_2 x$ is the inverse of the function $g(x)=2^x$, the graphs of these functions are reflections over the line $y = x$. The figure below shows the graphs of these two functions:



We can see that the functions are inverses by looking at the graph. For example, the graph of $g(x)=2^x$ contains the point (1, 2), while the graph of $f(x)=\log_2 x$ contains the point (2, 1).

Also, note that while that the graph of $g(x)=2^x$ is asymptotic to the x-axis, the graph of $f(x)=\log_2 x$ is asymptotic to the y-axis. This behavior of the graphs gives us a visual interpretation of the restricted range of g and the restricted domain of f.

When graphing log functions, it is important to consider x- values across the domain of the function. In particular, we should look at the behavior of the graph as it gets closer and closer to the asymptote. Consider $f(x)=\log_2 x$ for values of x between 0 and 1.

If $x=1/2$, then $f(1/2)=\log_2(1/2)=-1$ because $2^{-1}=1/2$
If $x=1/4$, then $f(1/4)=\log\_2(1/4)=-2$ because $2^{-2}=1/4$
If $x=1/8\}$, then $f(1/8)=\log_2(1/8)=-3$ because $2^{-3}=1/8$

From these values you can see that if we choose x values that are closer and closer to 0, the y values decrease (heading towards $-\infty$!). In terms of the graph, these values show us that the graph gets closer and closer to the y-axis. Formally we say that the vertical asymptote of the graph is $x = 0$.

Example 4: Graph the function $f(x)=\log_4 x$ and state the domain and range of the function.

Solution: The function $f(x)=\log_4 x$ is the inverse of the function $g(x) = 4^x$. We can sketch a graph of $f(x)$ by evaluating the function for several values of x, or by reflecting the graph of g over the line $y = x$.

If we choose to plot points, it is helpful to organize the points in a table:

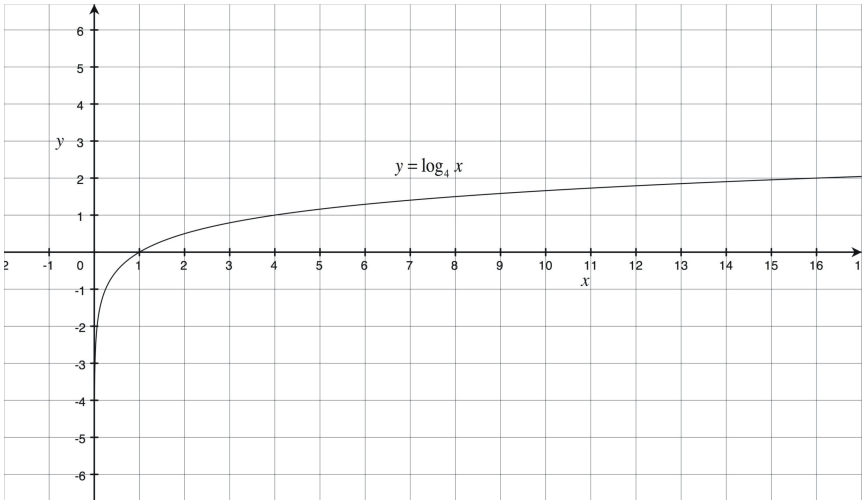
Table 4.16:

x	$y=\log_4 x$
1/4	

Table 4.16: (continued)

x	y=$\log_4 x$
1	0
4	1
16	2

The graph is asymptotic to the y-axis, so the domain of f is the set of all real numbers that are greater than 0. We can write this as a set:$\{x \in |x > 0\}$. While the graph might look as if it has a horizontal asymptote, it does in fact continue to rise. The range is .



A note about graphing calculators: You can use a graphing calculator to graph logarithmic functions, but many calculators will only allow you to use base 10 or base e. However, after the next lesson you will be able to rewrite *any* log as a log with base 10 or base e.

In this section we have looked at graphs of logarithmic functions of the form $f(x)$=$\log_b x$. Now we will consider the graphs of other forms of logarithmic equations.

Graphing Logarithmic Functions Using Transformations

As you saw in the previous lesson, you can graph exponential functions by considering the relationships between equations. For example, you can use the graph of $f(x)$=2^x to sketch a graph of $g(x)$=2^x + 3. Every y value of $g(x)$ is the same as a y value of $f(x)$, plus 3. Therefore we can shift the graph of $f(x)$ up 3 units to obtain a graph of $g(x)$.

We can use the same relationships to efficiently graph log functions. Consider again the log function $f(x)$=$\log_2 x$. The table below summarizes how we can use the graph of this function to graph other related function.

Table 4.17:

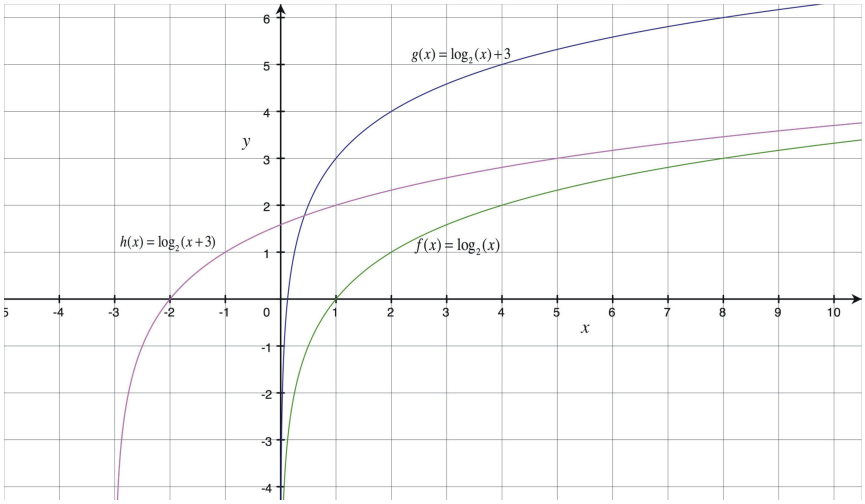
Equation	Relationship to $f(x)$=$\log_2 x$	Domain
$g(x)$=$\log_2(x$ - $a)$, for $a > 0$	Obtain a graph of g by shifting the graph of f a units to the right.	$x > a$
$g(x)$ = $\log_2(x$+$a)$ for $a > 0$	Obtain a graph of g by shifting the graph of f a units to the left.	$x >$ -a

Table 4.17: (continued)

Equation	Relationship to $f(x)=\log_2 x$	Domain
$g(x)=\log_2(x)$ + a for $a > 0$	Obtain a graph of g by shifting the graph of f up a units.	$x > 0$
$g(x)=\log_2(x)$-a for $a > 0$	Obtain a graph of g by shifting the graph of f down a units.	$x > 0$
$g(x)=a\log_2(x)$ for $a > 0$	Obtain a graph of g by vertically stretching the graph of f *by a factor of a.*	$x > 0$
$g(x)=-a\log_2(x)$, for $a > 0$	Obtain a graph of g by vertically stretching the graph of f *by a factor of a*, and by reflecting the graph over the x-axis.	$x > 0$
$g(x)=\log_2(-x)$	Obtain a graph of g by reflecting the graph of f over the y-axis.	$x < 0$

Example 5: Graph the functions $f(x)=\log_2(x)$,$g(x) = \log_2(x) + 3$, and $h(x) = \log_2(x + 3)$

Solution: The graph below shows these three functions together:



Notice that the location of the *3* in the equation makes a difference! When the 3 is added to $\log_2 x$, the shift is vertical. When the 3 is added to the x, the shift is horizontal. It is also important to remember that adding 3 to the x is a horizontal shift to the left. This makes sense if you consider the function value when x = -3:

$h(-3)=\log_2(-3 + 3)=\log_2 0$ = undefined

This is the vertical asymptote! To graph these functions, we evaluated them for certain values of x. But what if we want to know what the x value is for a particular y value? This means that we need to solve a logarithmic equation.

Solving Logarithmic Equations

In general, to solve an equation means to find the value(s) of the variable that makes the equation a true statement. To solve log equations, we have to think about what "log" *means*.

Consider the equation $\log_2 x=5$. What is the exponential form of this equation?
The equation $\log_2 x=5$ means that $2^5 = x$. So the solution to the equation is $x = 2^5 = 32$.

We can use this strategy to solve many logarithmic equations.

Example 6: Solve each equation for x:

Table 4.18:

a. $\log_4 x = 3$	b. $\log_5(x + 1) = 2$	c. $1 + 2\log_3(x - 5) = 7$

Solution: a. Writing the equation in exponential form gives us the solution: $x = 4^3 = 64$.

b. Writing the equation in exponential form gives us a new equation: $5^2 = x + 1$.
We can solve this equation for x:

Table 4.19:

5^2	$= x + 1$
25	$= x + 1$
x	$= 24$

c. First we have to isolate the log expression:

Table 4.20:

$1 + 2\log_3(x-5)$	$= 7$
$2\log_3(x - 5)$	$= 6$
$\log_3(x-5)$	$= 3$

Now we can solve the equation by rewriting it in exponential form:

Table 4.21:

$\log_3(x-5)$	$= 3$
3^3	$= x - 5$
27	$= x - 5$
x	$= 32$

We can also solve equations in which both sides of the equation contain logs. For example, consider the equation $\log_2(3x\text{-}1)=\log_2(5x - 7)$. Because the logarithms have the same base (2), the arguments of the log (the expressions $3x$ - 1 and $5x$ - 7) *must be equal*. So we can solve as follows:

Table 4.22:

	$\log_3(3x\text{-}1)$	$= \log_2(5x - 7)$
	$3x$ - 1	$= 5x$ - 7
	$+7$ $\vert+7$	
	$3x + 6$	$= 5x$
	$-3x$	
	6	$= 2x$
	x	$= 3$

Example 7: Solve for x: $\log_2(9x)=\log_2(3x + 8)$

Solution: The log equation implies that the expressions $9x$ and $3x + 8$ are equal:

Table 4.23:

	$log_2(9x)$	$= log_2(3x + 8)$
	$9x$	$= 3x + 8$
	$-3x$	$-3x$
	$6x$	$= 8$
	x	$= \frac{8}{6}$
	x	$= \frac{4}{3}$

Lesson Summary

In this lesson we have defined the logarithmic function as the inverse of the exponential function. When working with logarithms, it helps to keep in mind that a *logarithm is an exponent*. For example, $3 = \log_2 8$ and $2^3 = 8$ are two forms of the same numerical relationship among the three numbers 2, 3, and 8. The 2 is the base, the 3 is the exponent, and 8 is the 3rd power of 2.

Because logarithmic functions are the inverses of exponential functions, we can use our knowledge of exponential functions to graph logarithmic functions. You can graph a log function either by reflecting an exponential function over the line $y = x$, or by evaluating the function and plotting points. In this lesson you learned how to graph "parent" graphs such as $y = \log_2 x$ and $y = \log_4 x$, as well as how to use these parent graphs to graph more complicated log functions. When graphing, it is important to keep in mind that logarithmic functions have restricted domains. Each graph will have a vertical asymptote.

We can also use our knowledge of exponential relationships to solve logarithmic equations. In this lesson we solved 2 kinds of logarithmic equations. First, we solved equations by rewriting the equations in exponential form. Second, we solved equations in which both sides of the equation contained a log. To solve these equations, we used the following rule:

$\log_b f(x)=\log_b g(x) \rightarrow f(x) = g(x)$.

Points to Consider

1. What methods can you use to graph logarithmic functions?
2. What methods can you use to solve logarithmic equations?
3. What forms of log equations can you solve using the methods in this lesson? Can you write an equation that cannot be solved using these methods?

Review Questions

Write the exponential statement in logarithmic form.

1. $3^2 = 9$
2. $z^4 = 10$
3. Write the logarithmic statement in exponential form.
4. $\log_5 25 = 2$
5. $log_4 \frac{1}{6} = -1$
6. Complete the table of values for the function $f(x) = \log_3 x$

Table 4.24:

x	y = f(x)
1/9	
1/3	
1	
3	
9	

7.
8. Use the table above to graph $f(x)=\log_3 x$. State the domain and range of the function.
9. Consider $g(x) = -\log_3(x - 2)$a. How is the graph of $g(x)$ related to the graph of $f(x) = \log_3 x$?
   b. Graph $g(x)$ by transforming the graph of $f(x)$.
10. Solve each logarithmic equation: a. $\log_3 9x = 4$
    b. $7 + \log_2 x = 11$ (Hint: subtract 7 from both sides first.)
11. Solve each logarithmic equation:

Table 4.25:

a. $\log_5 6x = -1$	b. $\log_5 6x = \log_5(2x + 16)$	c. $\log_5 6x = \log_5(3x - 10)$

12.
13. Explain why the equation in 9c has no solution.

Review Answer

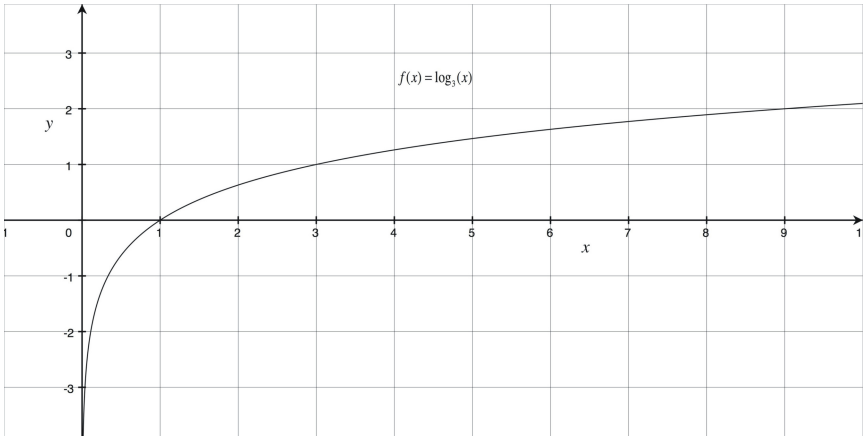
1. $3^2 = 9$
2. $z^4 = 10$
3. $5^2 = 25$
4. $6^{-1} = \frac{1}{6}$

5.

Table 4.26:

x	y = f(x)
1/9	
1/3	
1	0
3	1
9	2

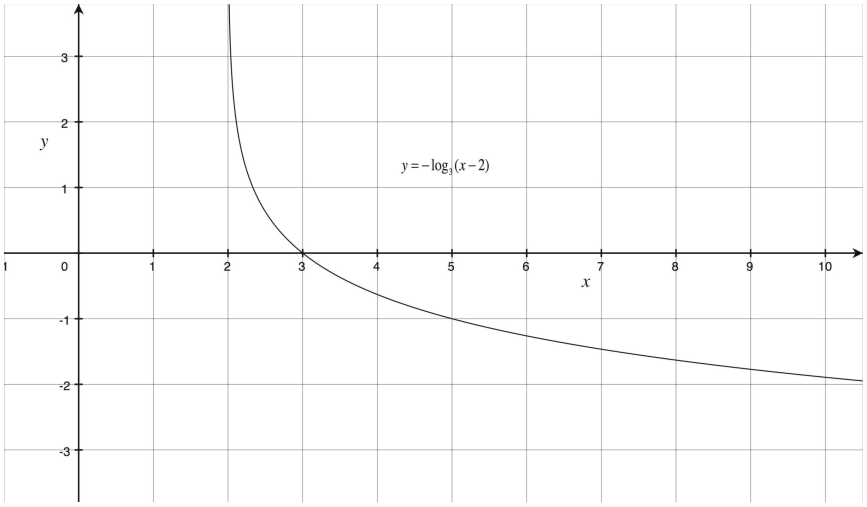
6.

7. 

D: All real numbers > 0

R: All real numbers.

8. a. The graph of g(x) can be obtained by shifting the graph of f(x) 2 units to the right, and reflecting it over the x- axis.b.



9. The solutions are:

Table 4.27:

a. x = 9	b. x = 16

10.

11. The solutions are:

Table 4.28:

a. x = 1/30	b. x = 4	c. no solution

12.
13. When we solve $6x$=$3x$-10 we find that x=-10/3, a value outside of the domain. Because there is no other x value that satisfies the equation, there is no solution.

Vocabulary

Argument The expression "inside" a logarithmic expression. The argument represents the "power" in the exponential relationship.

Asymptote An asymptote is line whose distance to a given curve tends to zero. An asymptote may or may not intersect its associated curve.

Domain The domain of a function is the set of all values of the independent variable (x) for which the function is defined.

Evaluate To evaluate a function is to identify a function value (y) for a given value of the independent variable (x).

Function A function is a relation between a domain (set of x values) and range (set of y values) in which every element of the domain is paired with one and only one element of the range. A function that is "one to one" is a function in which every element of the domain is paired with exactly one y value.

Logarithm The exponent of the power to which a base number must be raised to equal a given number.

Range The range of a function is the set of all function values, or values of the dependent variable (y).

4.4 Properties of Logarithms

Learning objectives

- Use properties of logarithms to write logarithmic expressions in different forms.
- Evaluate common logarithms and natural logarithms.
- Use the change of base formula and a scientific calculator to find the values of logs with any bases.

Introduction

In the previous lesson we defined the logarithmic function as the inverse of an exponential function, and we evaluated log expressions in order to identify values of these functions. In this lesson we will work with more complicated log expressions. We will develop properties of logs that we can use to write a log expression as the sum or difference of several expressions, or to write several expressions as a single log expression. We will also work with logs with base 10 and base e, which are the bases most often used in applications of logarithmic functions.

Properties of Logarithms

Because a logarithm is an exponent, the properties of logs are the same as the properties of exponents. Here we will prove several important properties of logarithms.

Property 1: $\log_b(xy) = \log_b x + \log_b y$

Proof: Let $\log_b x = n$ and $\log_b y = m$.

Rewrite both log expressions in exponential form:
$\log_b x = n \rightarrow b^n = x$
$\log_b y = m \rightarrow b^m = y$
Now multiply x and y: $xy = b^n \times b^m = b^{n+m}$
Therefore we have an exponential statement: $b^{n+m} = xy$.
The log form of the statement is: $\log_b xy = n + m$.
Now recall how we defined n and m:
$\log_b xy = n + m = \log_b x + \log_b y$.

Property 2: $\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y$

We can prove property 2 analogously to the way we proved property 1.

Proof: Let $\log_b x = n$ and $\log_b y = m$.

Rewrite both log expressions in exponential form:
$\log_b x = n \rightarrow b^n = x$
$\log_b y = m \rightarrow b^m = y$
Now divide x by y: $\frac{x}{y} = \frac{b^n}{b^m} = b^{n-m}$
Therefore we have an exponential statement: $b^{n-m} = \frac{x}{y}$.
The log form of the statement is: $\log_b\left(\frac{x}{y}\right) = n - m$.
Now recall how we defined n and m:
$\log_b\left(\frac{x}{y}\right) = n - m = \log_b x - \log_b y$.

Property 3: $\log_b x^n = n \log_b x$

The proof of the third property relies on another property of logs that we can derive by thinking about the definition of a log. Consider the expression $\log_2 2^{13}$. What does this expression *mean?*

The exponential form of $\log_2 2^{13} = ?$ is $2^? = 2^{13}$. Looking at the exponential form should convince you that the missing exponent is 13. That is, $\log_2 2^{13} = 13$. In general, $\log_b b^n = n$. This property will be used in the proof of property 3.

Proof (of Property 3):

Let $\log_b x = w$.
The exponential form of this log statement is $b^w = x$.
If we raise both sides of this equation to the power of n, we have $(b^w)^n = x^n$.

Using the power property of exponents, this equation simplifies to $b^{wn} = x^n$
If two expressions are equal, then the logs of both expressions are equal:

$\log_b b^{wn} = \log_b x^n$
Now consider the value of the left side of the equation: $\log_b b^{wn} = wn$.

Above, we showed that $b^{wn} = x^n$. By substitution, we have $\log_b x^n = wn$.
Above, we defined w: $\log_b x = w$. By substitution, we have
$\log_b x^n = (\log_b x)\ n = n\log_b x$.

We can use these properties to rewrite log expressions.

Expanding expressions

Using the properties we have derived above, we can write a log expression as the sum or difference of simpler expressions. Consider the following examples:

1. $\log_2 8x = \log_2 8 + \log_2 x = 3 + \log_2 x$
2. $log_3\left(\frac{x^2}{3}\right) = log_3 x^2 - log_3 3 = 2log_3 x - 1$

Using the log properties in this way is often referred to as "expanding". In the first example, expanding the log allowed us to simplify, as $\log 2\ 8 = 3$. Similarly, in the second example, we simplified using the log properties, and the fact that $\log_3 3 = 1$.

Example 1: Expand each expression:

Table 4.29:

a. $log_5 25x^2y$	b. $log_{10}\left(\frac{100x}{9b}\right)$

Solution:

a. $\log_5 25x^2\ y = \log_5 25 + \log_5 x^2 + \log_5 y = 2 + 2\log_5 x + \log_5 y$
b.

Table 4.30:

$log_{10}\left(\frac{100x}{9b}\right)$	$= log_{10}100x - log_{10}9b$
	$= log_{10}100 + log_{10}x - [log_{10}9 + log_{10}b]$ $= 2 + log_{10}x - log_{10}9 - log_{10}b$

Just as we can write a single log expression as a sum and difference of expressions, we can also write expanded expressions as a single expression.

Condensing expressions

To condense a log expression, we will use the same properties we used to expand expressions. Consider the expression $\log_6 8 + \log_6 27$. Alone, each of these expressions does not have an integer value. The value of $\log_6 8$ is between 1 and 2; the value of $\log_6 27$ is also between 1 and 2. If we condense the expression, we get:

$\log_6 8 + \log_6 27 = \log_6 (8 \times 27) = \log_6 216 = 3$

We can also condense algebraic expressions. This will be useful later for solving logarithmic equations.

Example 2: Condense each expression:

Table 4.31:

a. $2\log_3 x + \log_3 5x$ - $\log_3 (x + 1)$	b. $\log_2 (x^2 - 4)$ - $\log_2 (x + 2)$

Solution:

a. $2log_3x + log_35x - log_3(x + 1) = log_3x_2 + log_35x - log_3(x + 1)$

$= log_3(x^2(5x)) - log_3(x + 1)$
$= log_3\left(\frac{5x^3}{x+1}\right)$

b. $log_2(x^2 - 4) - log_2(x + 2) = log_2\left(\frac{x^2-4}{x+2}\right)$

$= log_2\left(\frac{(x+2)(x-2)}{x+2}\right)$
$= log_2(x - 2)$

It is important to keep in mind that a log expression may not be defined for certain values of x. First, the argument of the log must be positive. For example, the expressions in example 2b above are not defined for x 2 (which allows us to "cancel" $(x+2)$ without worrying about the condition x -2).

Second, the argument must be defined. For example, in example 2a above, the expression $\left(\frac{5x^3}{x+1}\right)$ is undefined if x = -1.

The log properties apply to logs with any real base. Next we will examine logs with base 10 and base e, which are the most common bases for logs (though only one is actually called "common").

Common logarithms and natural logs

A **common logarithm** is a log with base 10. We can evaluate a common log just as we evaluate any other log. A common log is usually written without a base.

Example 3: Evaluate each log

Table 4.32:

a. log 1	b. log 10	c. $log\sqrt{10}$

Solution:

a. log 1 = 0, as $10^0 = 1$.

b. log 10 = 1, as $10^1 = 10$
c. $log\ \sqrt{10} = \frac{1}{2}$ because $\sqrt{10} = 10^{1/2}$

As noted in lesson 3, logarithms were introduced in order to simplify calculations. After Napier introduced the logarithm, another mathematician, Henry Briggs, proposed that the base of a logarithm be standardized as 10. Just as Napier had labored to compile tables of log values (though his version of the logarithm is somewhat different from what we use today), Briggs was the first person to publish a table of common logs. This was in 1617!

Until recently, tables of common logs were included in the back of math textbooks. Publishers discontinued this practice when scientific calculators became readily available. A scientific calculator will calculate the value of a common log to 8 or 9 digits. Most calculators have a button that says LOG. For example, if you have TI graphing calculator, you can simply press LOG, and then a number, and the calculator will give you a log value up to 8 or 9 decimal places. For example, if you enter LOG(7), the calculator returns .84509804. This means that $10^{.84509804}$ 7. If we want to judge the reasonableness of this value, we need to think about powers of 10. Because $10^1 = 10$, log(7) should be less than 1.

Example 4: For each log value, determine two integers between which the log value should lie. Then use a calculator to find the value of the log.

Table 4.33:

a. log 50	b. log 818

Solution:

a. log 50
The value of this log should be between 1 and 2, as $10^1 = 10$, and $10^2 = 100$.
Using a calculator, you should find that log 50 1.698970004.
b. log 818
The value of this log should be between 2 and 3, as $10^2 = 100$, and $10^3 = 1000$.
Using a calculator, you should find that log 818 2.912753304.

The calculator's ability to produce log values is an example of the huge benefit that technology can provide. Only a few years ago, the calculations in the previous example would have each taken several minutes, while now they only take several seconds. While most people might not calculate log values in their every day lives, scientists and engineers are grateful to have such tools to make their work faster and more efficient.

Along with the LOG key on your calculator, you will find another logarithm key that says LN. This is the abbreviation for the **natural log**, the log with base e. Natural logs are written using "ln" instead of "log." That is, we write the expression $\log_e x$ as In x. How you evaluate a natural log depends on the argument of the log. You can evaluate some natural log expressions without a calculator. For example, ln $e = 1$, as $e^1 = e$. To evaluate other natural log expressions requires a calculator. Consider for example ln 7. Recall that e 2.7. This tells us that ln 7 should be slightly less than 2, as $(2.7)^2 = 7.29$. Using a calculator, you should find that ln 7 1.945910149.

Example 5: Find the value of each natural log.

Table 4.34:

a. $ln 100$	b. $ln\ \sqrt{e}$

Solution:

a. ln 100 is between 4 and 5. You can estimate this by considering powers of 2.7, or powers of 3: $3^4 = 81$, and $3^5 = 243$.
Using a calculator, you should find that ln 100 4.605171086.
b. Recall that a square root is the same as an exponent of 1/2. Therefore $ln\sqrt{e} = ln(e^{1/2}) = 1/2$

You may have noticed that the common log and the natural log are the only log buttons on your calculator. We can use either the common log or the natural log to find the values of logs with other bases.

Change of Base

Consider the log expression $\log_3 35$. The value of this expression is approximately 3 because $3^3 = 27$. In order to find a more exact value of $\log_3 35$, we can rewrite this expression in terms of a common log or natural log. Then we can use a calculator.

Let's consider a general log expression, $\log_b x = y$. This means that $b^y = x$. Recall that if two expressions are equal, then the logs of the expressions are equal. We can use this fact, and the power property of logs, to write $\log_b x$ in terms of common logs.

Table 4.35:

$b^y = x \Rightarrow logb^y = logx$	The logs of the expressions are equal

Table 4.36:

$\Rightarrow ylogb = logx$	Use the power property of logs
$\Rightarrow y = \frac{logx}{logb}$	Divide both sides by $logb$
$\Rightarrow log_b x = \frac{logx}{logb}$	Substitute $log_b x = y$

The final equation, $log_b x = \frac{logx}{logb}$, is called the change of base formula. Notice that the proof did not rely on the fact that the base of the log is 10. We could have used a natural log. Thus another form of the change of base formula is $\log_b x = \frac{lnx}{lnb}$.

Note that we could have used a log with any base, but we use the common log and the natural log so that we can use a calculator to find the value of an expression. Consider again $\log_3 35$. If we use the change of base formula, and then a calculator, we find that

$$log_3 35 = \frac{log35}{log3} = 3.23621727.$$

Example 6: Estimate the value, and then use the change of base formula to find the value of $\log_2 17$.

Solution: $\log_2 17$ is close to 4 because $2^4 = 16$ and $2^5 = 32$. Using the change of base formula, we have $log_2 17 = \frac{log17}{log2}$. Using a calculator, you should find that the approximate value of this expression is 4.087462841.

Lesson Summary

In this lesson we have developed and used properties of logarithms, including a formula that allows us to calculate the value of a log expression with any base. Out of context, it may seem difficult to understand the value of these kind of calculations. However, as you will see in later lessons in this chapter, we can use exponential and logarithmic functions to model a variety of phenomena.

Points to Consider

1. Why is the common log called common? Why 10?
2. Why would you want to estimate the value of a log before using a calculator to find its exact value?
3. What kind of situations might be modeled with a logarithmic function?

Review Questions

1. Expand the expression: $\log_b 5x^2$
2. Expand the expression: $\log_3 81x^5$
3. Condense the expression: $\log(x + 1) + \log(x - 1)$
4. Condense the expression: $3\ln(x) + 2\ln(y) - \ln(5x - 2)$
5. Evaluate the expressions: a. log 1000
   b. log 0.01
6. Evaluate the expressions: a. ln e^4
   b. $ln\left(\frac{1}{e^9}\right)$
7. Use the change of base formula to find the value of $\log_5 100$.
8. What is the difference between $\log_b x^n$ and $(\log_b x)^n$?
9. Condense the expression in order to simplify: $3 \log 2 + \log 125$
10. Is this equation true for any values of x and y? $\log_2 (x + y) = \log_2 x + \log_2 y$If so, give the values. If not, explain why not.

Review Answers

1. $\log_b 5 + 2 \log_b x$
2. $4 + 5 \log_3 x$
3. $\log (x^2 - 1)$
4. $ln\left(\frac{x^3y^2}{5x-2}\right)$
5. a. 3 b. -2
6. a. 4 b. -9
7. $\frac{log100}{log5} \approx 2.86$
8. The first expression is equivalent to $n \log_b x$. The second expression is the n^{th} power of the log.
9. $\log 1000 = 3$
10. $\log_2 (x + y) = \log_2 x + \log_2 y$ if and only if $x + y = xy$. The solutions to this equation are the possible values of x and y. For example, $x = 3$ and $y = 1.5$

Vocabulary

Common logarithm A common logarithm is a log with base $10k$. The log is usually written without the base.

Natural logarithm A natural log is a log with base e. The natural log is written as ln.

Scientific calculator A scientific calculator is an electronic, handheld calculator that will do calculations beyond the four operations $(+, -, \times, \div)$, such as square roots and logarithms. Graphing calculators will do scientific operations, as well as graphing and equation solving operations.

4.5 Exponential and Logarithmic Models and Equations

Learning objectives

- Analyze data to determine if it can be represented by an exponential or logarithmic model.
- Use a graphing calculator to find an exponential or logarithmic model, and use a model to answer questions about a situation.
- Solve exponential and logarithmic equations using properties of exponents and logarithms.
- Find approximate solutions to equations using a graphing calculator.

Introduction

So far in this chapter we have evaluated exponential and logarithm expressions, and we have graphed exponential and logarithmic functions. In this lesson you will extend what you have learned in two ways. First, we will introduce the idea of modeling real phenomena with logarithmic and exponential functions. Second, we will solve logarithmic and exponential equations. While you have already solved some equations in previous lessons, now you will be able to solve more complicated equations. This lesson will provide you with further tools for the applications of logarithmic and exponential functions that will be the focus of the remainder of the chapter.

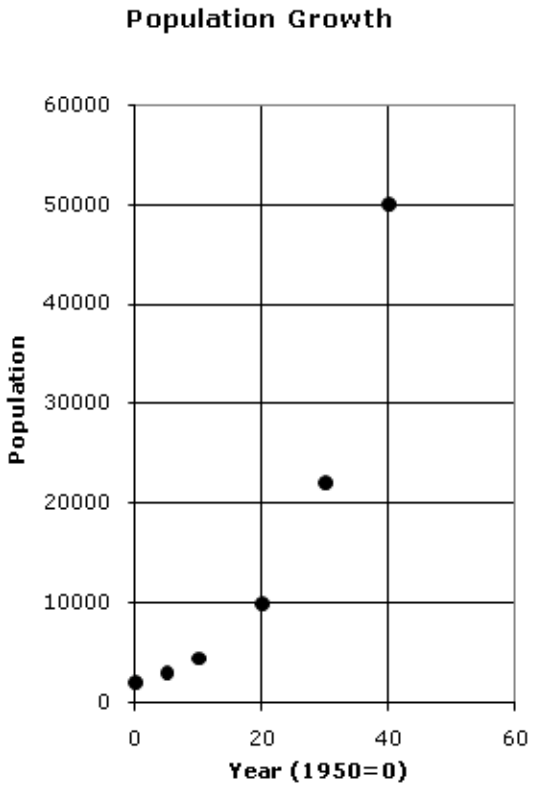
Exponential Models

Consider the following example: the population of a small town was 2,000 in the year 1950. The population increased over time, as shown by the values in the table:

Table 4.37:

Year (1950 = 0)	Population
0	2000
5	2980
10	4450
20	9900
30	22,000
40	50,000

If you plot these data points, you will see that the growth pattern is non-linear:



In many situations, population growth can be modeled with an exponential function (this is because population grows as a percentage of the current population, i.e. 8% per year). In lesson 7, you will learn how to create such models using information from a given situation. Here, we will focus on creating models using data and a graphing calculator.

The population data from the example above can be modeled with an exponential function, but the function is not unique. That is, there is more than one way to write a function to model this data. In the steps below you will see how to use a graphing calculator to find a function of the form $y = a(b^x)$ that fits the data in the table.

Technology Note

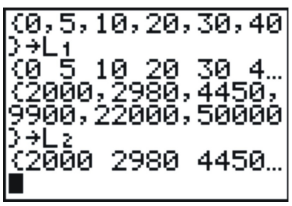
Using a TI-83/84 graphing calculator to find an exponential function that best fits a set of data

1. Entering the data

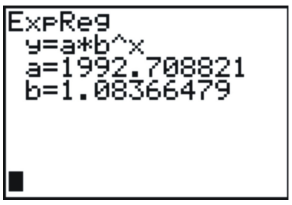
    a. Data must be entered into "lists". The calculator has six named lists, L1, L2, ... L6. We will enter the x values in L1 and the y values in L2. One way to do this is shown below:

Press <TI font_2nd> [{] and then enter the numbers separated by commas, and close by pressing the following: <TI font_2nd> [{]<TI font_STO> <TI font_2nd> [L1].

The top three lines of the figure below show the entry into list L1, followed by the entry of the y values into list L2.



Now press <TI font_STAT>, and move to the right to the CALC menu. Scroll down to option 10, **ExpReg.** Press <TI font_ENTER>, and you will return to the home screen. You should see **ExpReg** on the screen. As long as the numbers are in L1 and L2, the calculator will proceed to find an exponential function to fit the data you listed in List L1 and List L2. You should see on the home screen the values for a and b in the exponential function (See figure below). Therefore the function $y = 1992.7(1.0837)^x$ is an approximate model for the data.



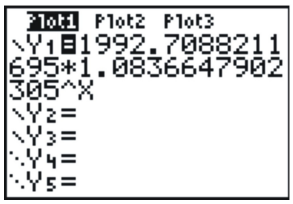
2. **Plotting the data and the equation**

To view plots of the data points and the equation on the same screen, do the following.

a. First, press <TI font_Y=> and clear any equations.

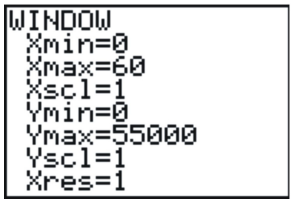
You can type in the equation above, or to get the equation from the calculator, do the following:

b. Enter the above rounded-off equation in Y1, or use the following procedure to get the full equation from the calculator: put the cursor in Y1, press <TI font_VARS>, 5, EQ, and 1. This should place the equation in Y1 (see figure below).

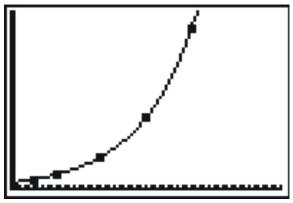


c. Now press <TI font_2nd>[STAT PLOT] and complete the items as shown in the figure below.

d. Now set your window. (Hint: use the range of the data to choose the window – the figure below shows our choices.)

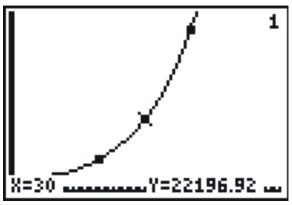


e. Press <TI font GRAPH> and you will to see the function and the data points as shown in the figure below.



3. Comparing the real data with the modeled results

It looks as if the data points lie on the function. However, using the TRACE function you can determine how close the modeled points are to the real data. Press <TI font_TRACE> to enter the TRACE mode. Then press the right arrow to move from one data point to another. Do this until you land on the point with value Y=22000. To see the corresponding modeled value, press the up or down arrow. See the figure below. The modeled value is approximately 22197, which is quite close to the actual data. You can verify any of the other data points using the same method.



Now that we have the equation $y = 1992.7(1.0837)^x$ to model the situation, we can estimate the population for any years that were not in the original data set. If we choose x values between 0 and 40, it is called **interpolation**. If we choose other x values outside of this domain, it is called **extrapolation**. Interpolation is, in a sense, a safer way of estimating population, because it is within the data points that we have, and does not require that we think about the end behavior of the function. For example, if we extrapolate to the year 1930, this means x = -20. The function value is 399. However, if the town was founded in 1940, then this data value does not make any sense. Similarly, if we extrapolate to the year 2000, we have x = 50. The function value is 110,711. However, if the town's pattern of population growth shifted (perhaps due to some economic change), this estimation could be highly inaccurate. As noted above, you will study exponential growth, as well as other exponential models, in the next two lessons. Now we turn to logarithmic models.

Logarithmic Models

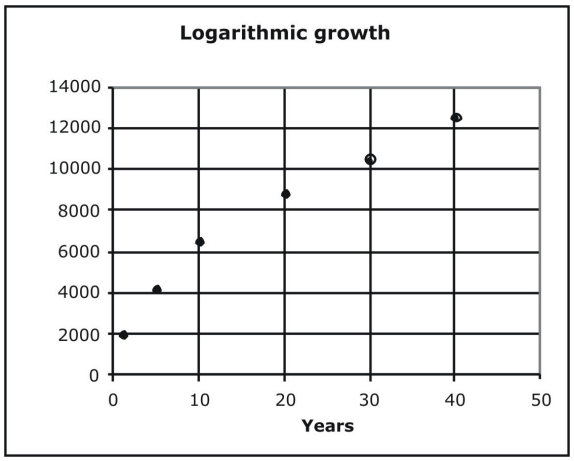
Consider another example of population growth:

Table 2

Table 4.38:

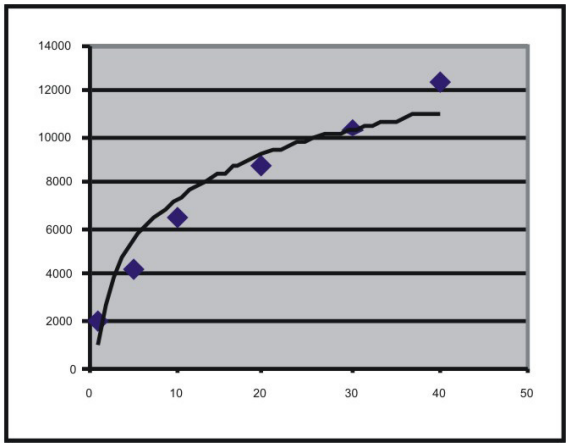
Year	Population
1	2000
5	4200
10	6500
20	8800
30	10500
40	12500

If we plot this data, we see that the growth is not quite linear, and it is not exponential either.



Just as we found an exponential model in the previous example, here we can find a logarithmic function to model this data. First enter the data in the table in L1 and L2. Then press STAT to get to the CALC menu. This time choose option 9. You should get the function $y = 930.4954615 + 2780.218173 \ln x$. If you view the graph and the data points together, as described in the Technology Note above, you will see that the graph of the function does not touch the data points, but models the general trend of the data.

Note about technology: you can also do this using an Excel spreadsheet. Enter the data in a worksheet, and create a scatterplot by inserting a chart. After you create the chart, from the chart menu, choose "add trendline." You will then be able to choose the type of function. Note that if you want to use a logarithmic function, the domain of your data set must be positive numbers. The chart menu will actually not allow you to choose a logarithmic trendline if your data include zero or negative x values. See below:



Solve Exponential Equations

Given an exponential model of some phenomena, such as population growth, you may want to determine a particular input value that would produce a given function value. Let's say that a function $P(x) = 2000(1.05)^x$ models the population growth for a town. What if we want to know when the population reaches 20,000?

To answer this equation, we must solve the equation $2000(1.05)^x = 20{,}000$. We can solve this equation by isolating the power $(1.05)^x$ and then using one of the log properties:

Table 4.39:

$2000(1.05)^x = 20,000$	Divide both sides of the equation by 2000
$(1.05)^x = 10$	Take the common log of both sides
$log(1.05)^x = log10$	Use the power property of logs
$xlog(1.05) = log10$	Evaluate $log10$
$xlog(1.05) = 1$	
$x = \frac{1}{log(1.05)} \approx 47$	
	Divide both sides by $log(1.05)$
	Use a calculator to estimate $log(1.05)$

We can use these same techniques to solve any exponential equation.

Example 1: Solve each exponential equation

Table 4.40:

a. $2^x + 7 = 19$	b. $3^{5x - 1} = 16$

Solution:

a. $2^x + 7 = 19$

$-7 - 7$
$2^x = 12$
$log2^x = log12$
$xlog2 = log12$

$x = \frac{log12}{log2} \approx 3.58$

b. $3^{5x-1} = 16$

$log3^{5x-1} = log16$
$(5x - 1)log3 = log16$

$5x - 1 = \frac{log16}{log3}$
$5x = \frac{log16}{log3} + 1$

$x = \frac{\frac{log16}{log3}+1}{5}$
$x \approx 0.705$

Solve Logarithmic Equations

In the previous lesson we solved two forms of log equations. Now we can solve more complicated equations, using our knowledge of log properties. For example, consider the equation $\log_2 (x) + \log_2 (x - 2) = 3$. We can solve this equation using a log property.

Table 4.41:

$\log_2 (x) + \log_2 (x - 2) = 3$	
$\log_2 (x(x - 2)) = 3$	$\log_b x + \log_b y = \log_b(xy)$
$\log_2 (x^2 - 2x) = 3 \Rightarrow$	write the equation in exponential form.
$2^3 = x^2 - 2x$	
$x^2 - 2x - 8 = 0$	Solve the resulting quadratic
$(x - 4)(x + 2) = 0$	
$x = -2, 4$	

The resulting quadratic has two solutions. However, only $x = 4$ is a solution to our original equataion, as $\log_2(-2)$ is undefined. We refer to $x = -2$ as an **extraneous solution.**

Example 2: Solve each equation

Table 4.42:

a. $\log (x + 2) + \log 3 = 2$	b. $\ln (x + 2) - \ln (x) = 1$

Solution:

a. $\log (x + 2) + \log 3 = 2$

Table 4.43:

$\log (3(x + 2)) = 2$	$\log_b x + \log_b y = \log_b (xy)$
$\log (3x + 6) = 2$	Simplify the expression $3(x+2)$
$10^2 = 3x + 6$	Write the log expression in exponential form
$100 = 3x + 6$	
$3x = 94$	Solve the linear equation
$x = 94/3$	

b. $\ln (x + 2) - \ln (x) = 1$

Table 4.44:

$ln\left(\frac{x+2}{x}\right) = 1$	$log_b x - log_b y = log_b\left(\frac{x}{y}\right)$
$e^1 = \frac{x+2}{x}$	Write the log expression in exponential form.
$ex = x + 2$	Multiply both sides by x.
$ex - x = 2$	Factor out x.
$x(e - 1) = 2$	Isolate x.
$x = \frac{2}{e-1}$	

The solution above is an *exact* solution. If we want a decimal approximation, we can use a calculator to find that $x \approx 1.16$. We can also use a graphing calculator to find an approximate solution, as we did in lesson 2 with exponential equations. Consider again the equation ln $(x + 2)$ - ln $(x) = 1$. We can solve this equation by solving a system:

$$\begin{cases} y = ln(x+2) - ln(x) \\ y = 1 \end{cases}$$

If you graph the system on your graphing calculator, as we did in lesson 2, you should see that the curve and the horizontal line intersection at one point. Using the **INTERSECT** function on the **CALC** menu (press <TI font_2nd>[CALC]), you should find that the x coordinate of the intersection point is approximately 1.16. This method will allow you to find approximate solutions for more complicated log equations.

Example 3. Use a graphing calculator to solve each equation:

Table 4.45:

a. $\log(5 - x) + 1 = \log x$	b. $\log_2(3x + 8) + 1 = \log_3 (10 - x)$

Solution:

a. $\log(5 - x) + 1 = \log x$

The graphs of $y = \log (5 - x) + 1$ and $y = \log x$ intersect at $x \approx 4.5454545$

Therefore the solution of the equation is $x \approx 4.54$.

b. $\log_2 (3x + 8) + 1 = \log_3 (10 - x)$

First, in order to graph the equations, you must rewrite them in terms of a common log or a natural log. The resulting equations are: $y = \frac{log(3x+8)}{log2} + 1$ and $y = \frac{log(10-x)}{log3}$. The graphs of these functions intersect at $x \approx -1.87$. This value is the approximate solution to the equation.

Lesson Summary

This lesson has introduced the idea of modeling a situation using an exponential or logarithmic function. When a population or other quantity has a steep increase over time, it may be modeled with an exponential function. When a population has a steep increase, but then slower growth, it may be modeled with a logarithmic function. (In a later lesson you will learn about a third option.) We have also examined techniques of solving exponential and logarithmic equations, based on our knowledge of properties of logarithms. The key property to remember is the power property:

$$\log_b x^n = n \log_b x$$

Using this property allows us to turn an exponential function into a linear function, which we can then solve in order to solve the original exponential function.

In the remaining lessons in this chapter, you will learn about several different real phenomena that are modeled with exponential and logarithmic equations. In these lessons you will also use the techniques of equation solving learned here in order to answer questions about these phenomena.

Points to Consider

1. What kinds of situations might be modeled with exponential functions or logarithmic functions?
2. What restrictions are there on the domain and range of data if we use these functions as models?
3. When might an exponential or logarithmic equation have no solution?
4. What are the advantages and disadvantages of using a graphing calculator to solve exponential and logarithmic equations?

Review Questions

For questions 1 - 5, solve each equation using algebraic methods. Give an exact solution.

1. $2\ (5^{x\ -\ 4}) + 7 = 43$
2. $4^{x} = 7^{3x\ -\ 5}$
3. $\log(5x + 200) + \log\ 2 = 3$
4. $\log_3\ (4x + 5)\ -\ \log_3\ x = 2$
5. $\ln\ (4x + 1)\ -\ \ln\ (2x) = 3$
6. Use a graphing utility to solve the equation in #4.
7. Use a graphing utility to solve the equation $\log(x^2\ -3) = \log\ (x + 5)$
8. In example 3b, the solution to the log equation $\log_2(3x + 8) + 1 = \log_3\ (10\ -\ x)$ was found to be $x\ -\ 1.87$. One student read this example, and wondered how the value of x could be negative, given that you cannot take a log of a negative number. How would you explain to this student why the solution is valid?
9. The data set below represents a hypothetical situation: You invest $2000 in a money market account, and you do not invest more money or withdraw any from the account. **Table 3**

Table 4.46:

Time since you invested (in years)	Amount in account
0	2,000
2	2200
5	2500
10	3300
20	4500

10.
11. a. Use a graphing utility to find an exponential model for the data. b. Use your model to estimate the value of the account after the 8
12. th
13. year. c. At this rate, much money would be in the account after 30 years? d. Explain how your estimate in part c might be inaccurate. (What might happen after 20 years?)
14. The data set below represents the growth of a plant. **Table 4**

Table 4.47:

Time since planting (days)	Height of plant (inches)
1	.2
4	.5
5	.57

Table 4.47: (continued)

Time since planting (days)	Height of plant (inches)
10	1.2
12	1.3
14	1.4

15.
16. a. Use a graphing utility to find a logarithmic equation to model the data. b. Use your model to estimate the height of the plant after 15 days. Compare this estimate to the trend in the data. c. Give an example of an
17. x
18. value for which the model does not make sense.
19. In the lesson, the equation log(5 - x) + 1 = log x was solved using a graph. Solve this equation algebraically in order to (a) verify the approximate solution found in the lesson and (b) give an exact solution.

Review Answers

1. $log_5 18 + 4$ or $\frac{log 18}{log 5} + 4$
2. $\frac{-5log 7}{log 4 - 3log 7}$
3. x = 60
4. x = 1
5. $x = \frac{1}{2e^3 - 4}$
6. The function y = $\log_3$ (4x + 5) - $\log_3$ x intersects the line y = 2 at the point (1, 2)
7. The graphs intersect twice, giving 2 solutions: x -2.37, x 3.37
8. The value of can be negative as long as the argument of the log is positive. In this equation, the arguments are 3x+8 and 10-x. Neither expression takes on a negative value for x -1.87
9. a. y = 2045.405(1.042)^{x}b. About $2840
   c. About $7003
   d. After that much time, you may decide to withdraw the money to spend or to invest in something with more potential for growth.
10. a. y = 0.0313 + .4780 ln xb. The model gives 1.32 inches. The data would suggest the plant is at least 1.4 inches tall.
    c. The model does not make sense for negative x values. Also, at some point the plant could die. This reality puts an upper bound on x.
11. log (5 - x) + 1 = log $x$log (5 - x) - log x + 1 = 0
    log(5 - x) - log x = -1
    $log\left(\frac{5-x}{x}\right) = -1$
    $10^{-1} = \frac{5-x}{x}$
    $0.1 = \frac{5-x}{x}$
    0.1 x = 5 - x
    1.1 x = 5
    $x = \frac{5}{1.1} = \frac{50}{11} = 4\frac{6}{11} = 5.\overline{54}$

Vocabulary

Extraneous solution An extraneous solution is a solution to an equation used to solve an initial equation that is not a solution to the initial equation. Extraneous solutions occur when solving certain kinds of equations, such as log equations, or square root equations.

Extrapolation To extrapolate from data is to create new data points, or to predict, outside of the domain of the data set.

Interpolation To interpolate is to create new data points, or to predict, within the domain of the data set, but for points not in the original data set.

4.6 Compound Interest

Learning objectives

- Calculate compound interest, including continuous compounding.
- Compare compound interest situations.
- Determine algebraically and graphically the time it takes for an account to reach a particular value.

Introduction

In the previous lesson, you learned about modeling growth using an exponential function. In this lesson we will focus on a specific example of exponential growth: compounding of interest. We will begin with the case of simple interest, which refers to interest that is based only on the principal, or initial amount of an investment or loan. Then we will look at what it means for interest to compound. In the simplest terms, compounding means that interest accrues (you gain interest on an investment, or owe more on a loan) based on the principal you invested, as well as on interest you have already accrued. As you will soon see, compound interest is a case of exponential growth. In this lesson we will look at specific examples of compound interest, and we will write equations to model these specific situations.

Simple interest over time

As noted above, simple interest means that interest accrues based on the principal of an investment or loan. The simple interest is calculated as a percent of the principal. The formula for simple interest is, in fact, simple:

The variable P represents the principal amount, r represents the interest rate, and t represents the amount of time the interest has been accruing. For example, say you borrow \$2,000 from a family member, and you insist on repaying with interest. You agree to pay 5% interest, and to pay the money back in 3 years. The interest you will owe will be $2000(0.05)(3) = \$300$. This means that when you repay your loan, you will pay \$2300. Note that the interest you pay after 3 years is not 5% of the original loan, but 15%, as you paid 5% of \$2000 each year for 3 years.

Now let's consider an example in which interest is compounded. Say that you invest \$2000 in a bank account, and it earns 5% interest annually. How much is in the account after 3 years?

In order to determine how much money is in the account after three years, we have to determine the amount of money in the account after each year. The table below shows the calculations for one, two, and three years of this investment:

Table 4.48:

Year	Principal + interest
After one year	$2000 + 2000(0.05) = 2000 + 100 = \$2100$
After 2 years	$2100 + 2100(0.05) = 2100 + 105 = \$2205$
After 3 years	$2205 + 2205(0.05) = 2205 + 110.25 = \$2315.25$

Therefore, after three years, you will have \$2315.25 in the account, which means that you will have earned \$315.25 in interest. With simple interest, you would have earned \$300 in interest. Compounding results in more interest because the principal on which the interest is calculated increased each year. For example, in the second year shown in the table above, you earned 5% of 2100, *not* 5% of 2000, as would be the case of simple interest. The main idea here is that compounding creates more interest because you are earning interest on interest, and not just on the principal.

But how much more?

You might look at the above example and say, "it's only \$15.25." Remember that we have only looked at one example, and this example is a small one: in the grand scheme of investing, \$2000 is a small amount of money, and we have only looked at the growth of the investment for a short period of time. For example, if you are saving for retirement, you could invest for a period of 30 years or more, and you might invest several thousand dollars each year.

The formulas and methods for calculating retirement investments are more complicated than what we will do here. However, we can use the above example to derive a formula that will allow us to calculate compound interest for any number of years.

The compound interest formula

To derive the formula for compound interest, we need to look at a more general example. Let's return to the previous example, but instead of assuming the investment is \$2000, let the principal of the investment be P dollars. The key idea is that each year you have 100% of the principal, plus 5% of the previous balance. The table below shows the calculations of this more general investment.

Table 4.49:

Year	Principal + interest	New principal
1	$P + P \times .05 = 1.00P + .05P =$	$1.05P$
2	$1.05P + .05 \times (1.05P) = 1.05P$ $[1+ .05] = 1.05P \times 1.05 =$	$(1.05)^2$ P
3	$(1.05)^2$ P + .05 × $(1.05)^2$ P = $(1.05)^2$ P [1 + .05] = $(1.05)^2$ P [1.05] =	$(1.05)^3$ P

Notice that at the end of every year, the amount of money in the investment is a power of 1.05, times P, and that the power corresponds to the number of years. Given this pattern, you might hypothesize that after 4 years, the amount of money is $(1.05)^4$ P.

We can generalize this pattern to a formula. As above, we let P represent the principal of the investment.

Now, let t represent the number of years, and r represent the interest rate. Keep in mind that $1.05 = 1+0.05$. So we can generalize:

$$A(t) = P(1 + r)^t$$

This function will allow us to calculate the amount of money in an investment, if the interest is compounded each year for t years.

Example 1: Use the formula above to determine the amount of money in an investment after 20 years, if you invest \$2000, and the interest rate is 5% compounded annually.

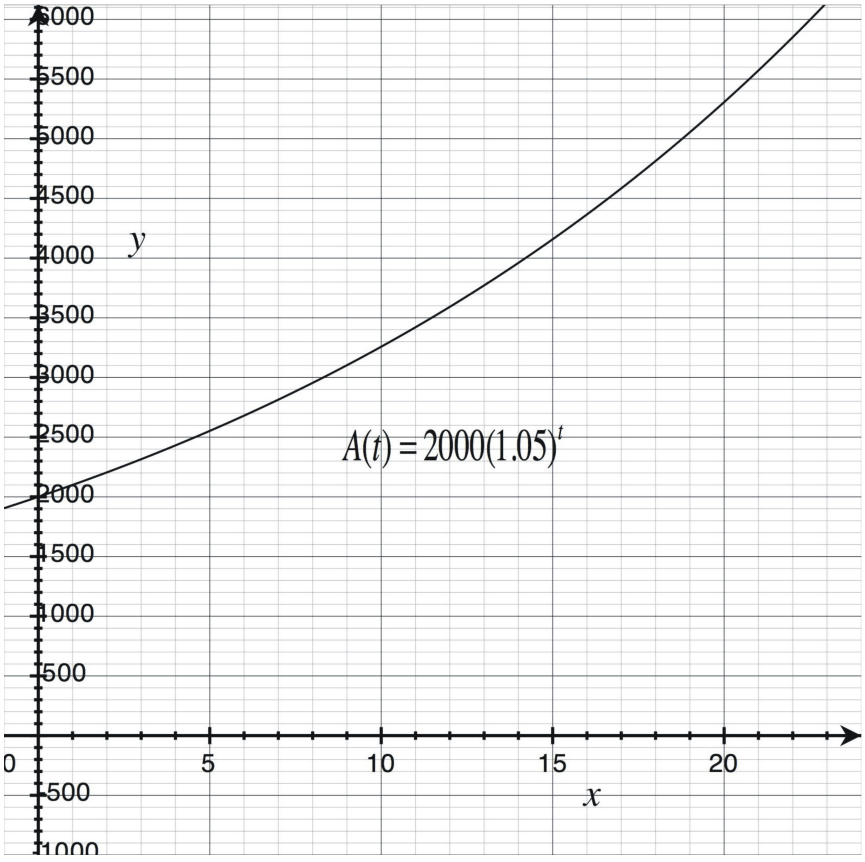
Solution: The investment will be worth \$5306.60

$A(t) = P(1 + r)^t$

$A(20) = 2000(1.05)^{20}$

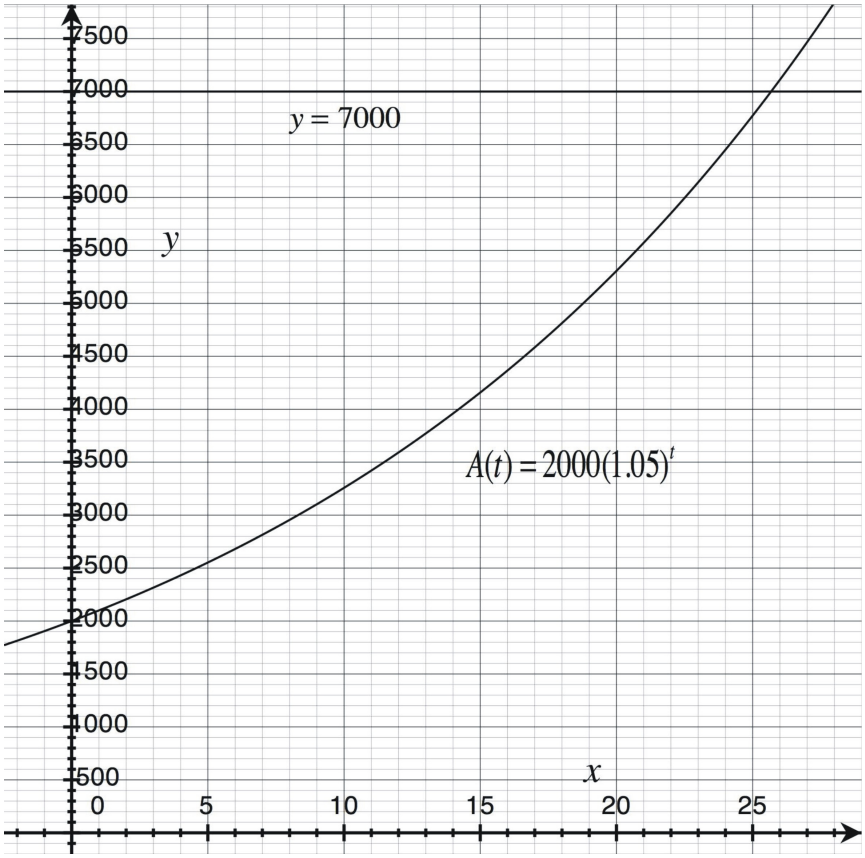
$A(20) = \$5306.60$

In the above example, we found the value of this investment after a particular number of years. If we graph the function $A(t) = 2000(1.05)^t$, we can see the values for any number of years.



If you graph this function using a graphing calculator, you can determine the value of the investment by tracing along the function, or by pressing <TI font_TRACE> on your graphing calculator and then entering an x value. You can also choose an investment value you would like to reach, and then determine the number of years it would take to reach that amount. For example, how long will it take for the investment to reach \$7,000?

As we did earlier in the chapter, we can find the intersection of the exponential function with the line $y = 7000$.



You can see from Figure 2 that the line and the curve intersect at a little less than $x = 26$. Therefore it would take almost 26 years for the investment to reach \$7000.

You can also solve for an exact value:

Table 4.50:

$2000(1.05)^t = 7000$	
$(1.05)^t = \frac{7000}{2000}$	Divide both sides by 2000 and simplify the right side of the equation.
$(1.05)^t = 3.5$	Take the ln of both sides (you can use any log, but ln or log base 10 will allow you to use a calculator.)
$ln(1.05)^t = ln3.5$	
$t[ln(1.05)] = ln3.5$	Use the power property of logs
$t = \frac{ln3.5}{ln1.05} \approx 25.68$	Divide both sides by ln 1.05

The examples we have seen so far are examples of annual compounding. In reality, interest is often compounded more frequently, for example, on a monthly basis. In this case, the interest rate is divided amongst the 12 months. The formula for calculating the balance of the account is then slightly different:

$$A(t) = P\left(1 + \frac{r}{12}\right)^{12t}$$

Notice that the interest rate is divided by 12 because $1/12^{\text{th}}$ of the rate is applied each month. The variable t in the exponent is multiplied by 12 because the interest is calculated 12 times per year.

In general, if interest is compounded n times per year, the formula is:

$$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$$

Example 2: Determine the value of each investment.

a. You invest \$5000 in an account that gives 6% interest, compounded monthly. How much money do you have after 10 years?
b. You invest \$10,000 in an account that gives 2.5% interest, compounded quarterly. How much money do you have after 10 years?

Solution:

a. \$5000, invested for 10 years at 6% interest, compounded monthly.

$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$

$A(10) = 5000\left(1 + \frac{.06}{12}\right)^{12 \cdot 10}$

$A(10) = 5000\,(1.005)^{120}$

$A(10) \approx 9096.98$

b. \$10000, invested for 10 years at 2.5% interest, compounded quarterly.

Quarterly compounding means that interest is compounded four times per year. So in the equation, $n = 4$.

$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$

$A(10) = 6000\left(1 + \frac{.025}{4}\right)^{4 \cdot 10}$

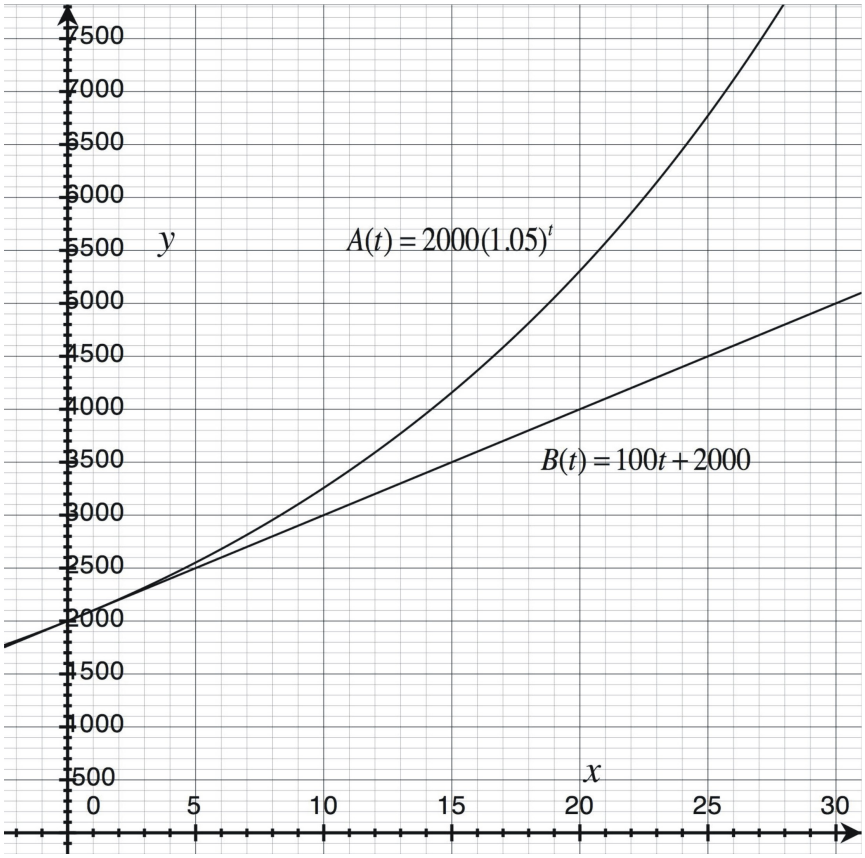
$A(10) = 6000(1.00625)^{40}$

$A(10) \approx 12{,}830.30$

In each example, the value of the investment after 10 years depends on three quantities: the principal of the investment, the number of compoundings per year, and the interest rate. Next we will look at an example of one investment, but we will vary each of these quantities.

The power of compound interest

Consider the investment in example 1: \$2000 was invested at an annual interest rate of 5%. We modeled this situation with the equation $A(t) = 2000(1.05)^{\text{t}}$. We can use this equation to determine the amount of money in the account after any number of years. As we saw above, the value of the account grows exponentially. You can see how fast the investment grows if we compare it to linear growth. For example, if the same investment earned simple interest, the value of the investment after t years could be modeled with the function $B(t) = 2000 + 2000(.05)t$. We can simplify this to be: $B(t) = 100t + 2000$. The exponential function and this linear function are shown here.



Notice that if we look at these investments over a long period of time (30 years are shown in the graph), the values look very close together for x values less than 10. For example, after 5 years, the compound interest investment is worth \$2552.60, and the simple interest investment is worth \$2500. But, after 20 years, the compound interest investment is worth \$5306.60, and the simple interest investment is worth \$4,000. After 20 years, simple interest has doubled the amount of money, while compound interest has resulted in 2.65 times the amount of money.

The main idea here is that an exponential function grows faster than a linear one, which you can see from the graphs above. But what happens to the investment if we change the interest rate, or the number of times we compound per year?

Example 3: Compare the values of the investments shown in the table. If everything else is held constant, how does the interest rate influence the value of the investment?

Table 4.51:

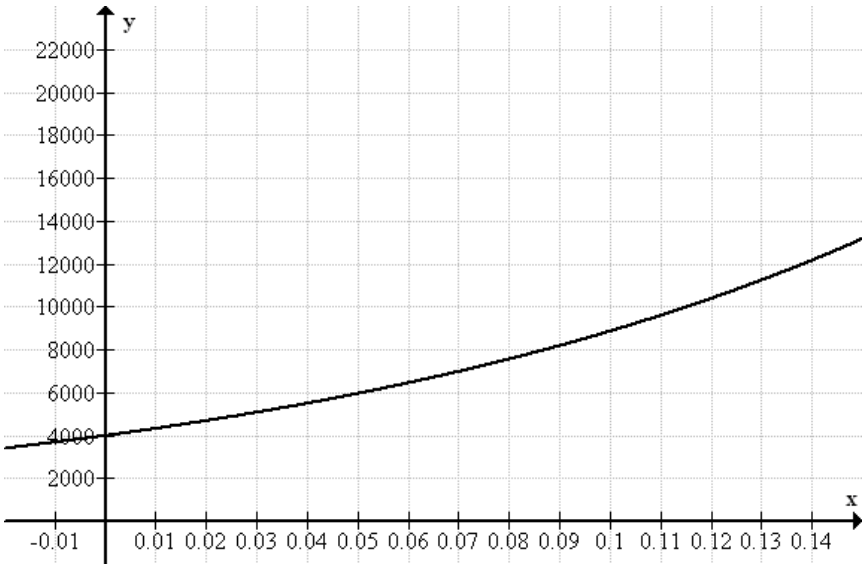
	Principal	r	n	t
a.	\$4,000	.02	12	8
b.	\$4,000	.05	12	8
c.	\$4,000	.10	12	8
d.	\$4,000	.15	12	8
e.	\$4,000	.22	12	8

Solution: Using the compound interest formula $A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$, we can calculate the value of each investment. In all cases, we have $A(8) = 4000\left(1 + \frac{r}{12}\right)^{12 \cdot 8}$.

Table 4.52:

	Principal	r	n	t	A
a.	\$4,000	.02	12	8	\$4693.42
b.	\$4,000	.05	12	8	\$5962.34
c.	\$4,000	.10	12	8	\$8872.70
	\$4,000	.15	12	8	\$13182.05
e.	\$4,000	.22	12	8	\$22882.11

As we increase the interest rate, the value of the investment increases. It is part of every day life to want to find the highest interest rate possible for a bank account (and the lowest possible rate for a loan!). Let's look at just how fast the value of the account grows. Remember that each calculation in the table above started with $A(8) = 4000\left(1 + \frac{r}{12}\right)^{12 \cdot 8}$. Notice that this is a function of r, the interest rate. We can write this equation in a more standard form: $f(x) = 4000\left(1 + \frac{x}{12}\right)^{96}$. The graph of this function is shown below:



Notice that while this function is not exponential, it does grow quite fast. As we increase the interest rate, the value of the account increases very quickly.

Example 4: Compare the values of the investments shown in the table. If everything else is held constant, how does the compounding influence the value of the investment?

Table 4.53:

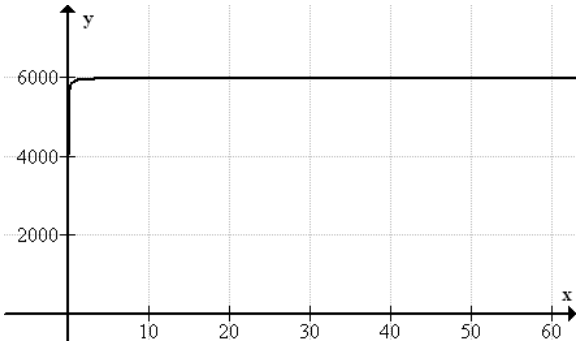
	Principal	r	n	t
a.	\$4,000	.05	1 (annual)	8
b.	\$4,000	.05	4 (quarterly)	8
c.	\$4,000	.05	12 (monthly)	8
d.	\$4,000	.05	365 (daily)	8
e.	\$4,000	.05	8760 (hourly)	8

Solution: Again, we use the compound interest formula. For this example, the n is the quantity that changes: $A(8) = 4000\left(1 + \frac{.05}{n}\right)^{8n}$

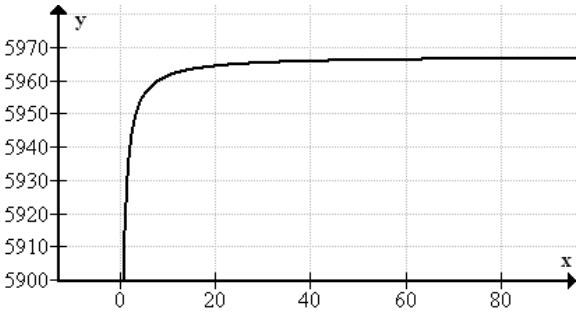
Table 4.54:

	Principal	r	n	t	A
a.	\$4,000	.05	1 (annual)	8	\$5909.82
b.	\$4,000	.05	4 (quarterly)	8	\$5952.52
c.	\$4,000	.05	12 (monthly)	8	\$5962.34
d.	\$4,000	.05	365 (daily)	8	\$5967.14
e.	\$4,000	.05	8760 (hourly)	8	\$5967.29

In contrast to the changing interest rate, in this example, increasing the number of compoundings per year does not seem to dramatically increase the value of the investment. We can see why this is the case of we look at the function $A(8) = 4000\left(1 + \frac{.05}{n}\right)^{8n}$. A graph of the function $f(x) = 4000\left(1 + \frac{.05}{x}\right)^{8x}$ is shown below:



The graph seems to indicate that the function has a horizontal asymptote at \$6000. However, if we zoom in, we can see that the horizontal asymptote is closer to 5967.



What does this mean? This means that for the investment of \$4000, at 5% interest, for 8 years, compounding more and more frequently will never result in more than about \$5968.00.

Another way to say this is that the function $f(x) = 4000\left(1 + \frac{.05}{x}\right)^{8x}$ has a limit as x approaches infinity. Next we will look at this kind of limit to define a special form of compounding.

Continuous compounding

Consider a hypothetical example: you invest \$1.00, at 100% interest, for 1 year. For this situation, the amount of money you have at the end of the year depends on how often the interest is compounded:

$$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$$

$A(t) = 1\left(1 + \frac{1}{n}\right)^{ln}$
$A = \left(1 + \frac{1}{n}\right)^{n}$

Now let's consider different compoundings:

Table 4.55:

Compounding	N	A ≈
Annual	1	2
Quarterly	4	2.44140625
Monthly	12	2.61303529022
Daily	365	2.71456748202
Hourly	8,760	2.71812669063
By the minute	525,600	2.7182792154
By the second	31,536,000	2.71828247254

The values of A in the table have a limit, which might look familiar: it's the number e. In fact, one of the definitions of e is $\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^{n}$.

A related limit is one that will lead us to a special kind of compound interest: $\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^{n} = e^{x}$. (The proof of this limit requires calculus. However, in one of the review questions, you will examine this limit more closely.)

Now we can define what is known as continuous compounding. If interest is compounded n times per year, the equation we use is: $A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$. We can also write the function as $A(t) = P\left((1 + \frac{r}{n})^{n}\right)^{t}$. If we compound more and more often, we are looking at what happens to this function as $n \to \infty$. Recall the limit above: $\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^{n} = e^{x}$. Here, this means $\lim_{n\to\infty}\left(1 + \frac{r}{n}\right)^{n} = e^{r}$. So as n approaches ∞, $\left((1 + \frac{r}{n})^{n}\right)^{t}$ approaches $P(e^{r})^{t} = Pe^{rt}$.

The function $A(t) = Pe^{rt}$ is the formula we use to calculate the amount of money when interest is continuously compounded, rather than interest that is compounded at discrete intervals, such as monthly or quarterly. For example, consider again the investment in example 1 above: what is the value of an investment after 20 years, if you invest \$2000, and the interest rate is 5% compounded continuously?

$A(t) = Pe^{\mathrm{rt}}$
$A(20) = 2000e^{.05(20)}$
$A(20) = 2000e^{1}$
$A(20) = \$5436.56$

Just as we did with the standard compound interest formula, we can also determine the time it takes for an account to reach a particular value if the interest is compounded continuously.

Example 5: How long will it take \$2000 to grow to \$25,000 in the previous example?

Solution: It will take about 50 years:

Table 4.56:

$A(t) = Pe^{rt}$	
$25{,}000 = 2000e^{.05(t)}$	
$12.5 = e^{.05(t)}$	Divide both sides by 2000
$\ln 12.5 = \ln e^{.05(t)}$	Take the ln of both sides
$\ln 12.5 = .05t \ln e$	Use the power property of logs
$\ln 12.5 = .05t \times 1$	$\ln e = 1$
$\ln 12.5 = 0.5t$	Isolate t
$t = \frac{ln12.5}{.05} \approx 50.5$	

Lesson Summary

In this lesson we have developed formulas to calculate the amount of money in a bank account or an investment when interest is compounded, either a discrete number of times per year, or compounded continuously. We have found the value of accounts or investments, and we have found the time it takes to reach a particular value. We have solved these problems algebraically and graphically, using our knowledge of functions in general, and logarithms in particular.

In general, the examples we have seen are conservative in the larger scheme of investing. Given all of the information available today about investments, you may look at the examples and think that the return on these investments seems low. For example, in the last example, 50 years probably seems like a long time to wait!

It is important to keep in mind that these calculations are based on an initial investment only. In reality, if you invest money long term, you will invest on a regular basis. For example, if an employer offers a retirement plan, you may invest a set amount of money from every paycheck, and your employer may contribute a set amount as well. As noted above, the calculations for the growth of a retirement investment are more complicated. However, the exponential functions you have studied in this lesson are the basis for the calculations you would need to do. The examples here are meant to illustrate an application of exponential functions, and the power of compound interest.

Points to Consider

1. Why is compound interest modeled with an exponential function?
2. What is the difference between compounding and continuous compounding?
3. How are logarithms useful in solving compound interest problems?

Review Questions

1. You put \$3500 in a bank account that earns 5.5% interest, compounded monthly. How much is in your account after 2 years? After 5 years?
2. You put \$2000 in a bank account that earns 7% interest, compounded quarterly. How much is in your account after 10 years?
3. Solve an exponential equation in order to answer the question: given the investment in question 2, how many years will it take for the account to reach \$10,000?
4. Use a graph to verify your answer to question 3.

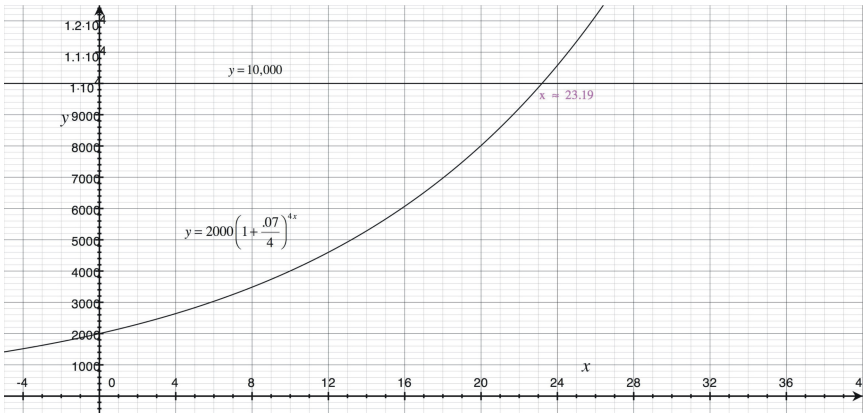
5. Consider two investments: (1) \$2000, invested at 6% interest, compounded monthly
   (2) \$3000, invested at 4.5% interest, compounded monthly
   Use a graph to determine when the 2 investments have equal value.
6. You invest \$3000 in an account that pays 6% interest, compounded monthly. How long does it take to double your investment?
7. Explain why the answer to #6 does not depend on the amount of the initial investment
8. You invest \$4,000 in an account that pays 3.2% interest, compounded continuously. What is the value of the account after 5 years?
9. You invest \$6,000 in an account that pays 5% interest, compounded continuously. What is the value of the account after 10 years?
10. Consider the investment in example 8. How many years will it take the investment to reach \$20,000?
11. In this lesson, we introduced this limit: $\lim n \to \infty \left(1 + \frac{r}{n}\right)^n = e^r$We noted in the lesson that the proof of this limit requires calculus. Here, we will examine a few specific cases in order to see how this limit is true.
    Using a graphing calculator, estimate the value of this limit for the given values of x. Do the limits seem to match the value of e^x? *Do these calculations suffice as a proof of the limit? Why or why not?*
    Hint: Graph each limit expression as a function, where x represents n.

Table 4.57:

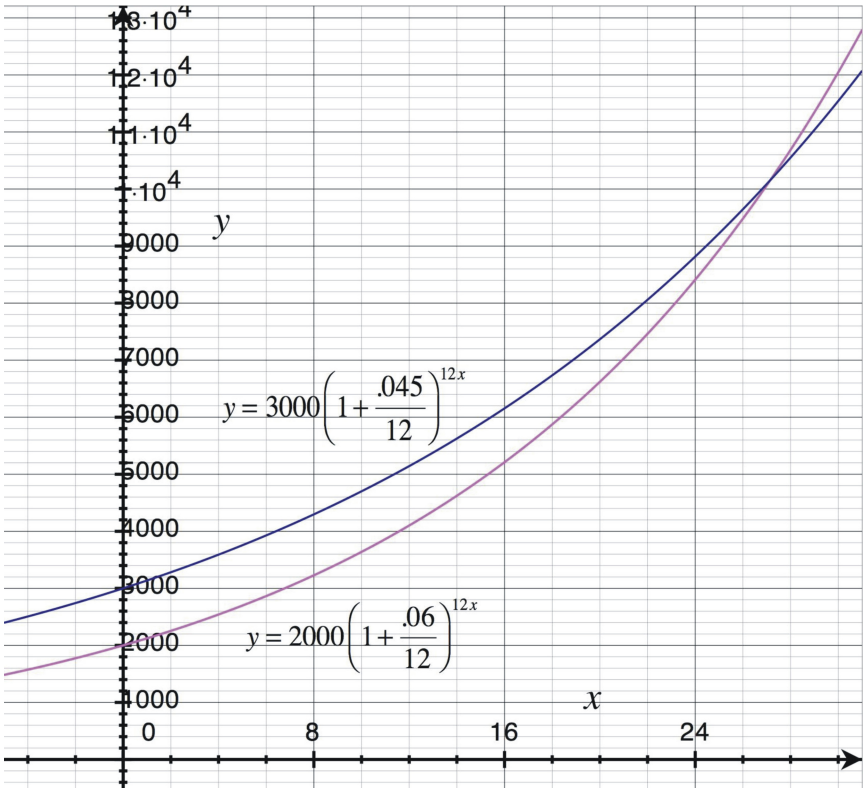
r	$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$	e^r
0		
1		
2		
3		
4		
5		
10		

12.

Review Answers

1. After 2 years: \$3905.99. After 5 years: \$4604.96.
2. \$4003.20
3. $t = \frac{ln5}{4ln1.0175} \approx 23.19$ years.
4. 

The functions cross at $x \approx 23.19$

5. 

It takes about 27 years for the two investments to have the same value.
6. $t = \frac{ln2}{12ln1.005} \approx 11.58$ years.
7. When solving for t, the 6000 is divided by 3000, resulting in a 2 on the left side of the equation. (Hence the ln 2.) This would be the same, no matter what the initial investment was.
8. $4694.03
9. $9892.36
10. It will take about 50 years.
11.

Table 4.58:

r	$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$	e^r
0	1	1
1	e	e
2	≈ 7.389	≈ 7.389
3	≈ 20.086	≈ 20.086
4	≈ 54.598	≈ 54.598
5	≈ 148.413	≈ 148.413
10	≈ 22026.466	≈ 22026.466

12.
13. The values in the table match, but this does not count as a proof. A proof needs to show that the values match for ALL values of
14. r
15. .

Vocabulary

Accrue To accrue is to increase in amount or value over time. If interest accrues on a bank account, you will have more money in your account. If interest accrues on a loan, you will owe more money to your lender.

Compound interest Compound interest is interest based on a principal and on previous interest earned.

Continuous compounding Interest that is based on continuous compounding is calculated according to the equation $A(t)=Pe^{rt}$, where P is the principal, r is the interest rate, t is the length of the investment, and A is the value of the account or investment after t years.

Principal The principal is the initial amount of an investment or a loan.

Simple interest Simple interest is interest that is calculated as a percent of the principal, as a function of time.

4.7 Growth and Decay

Learning objectives

- Model situations using exponential and logistic functions.
- Solve problems involving these models, using your knowledge of properties of logarithms, and using a graphing calculator.

Introduction

In lesson 5 you learned about modeling phenomena with exponential and logarithmic functions. In the examples in lesson 5, you used a graphing calculator to find a line that fits a given set of data. Here we will use algebraic techniques to develop models, and you will learn about another kind of function, the logistic function, that can be used to model growth.

Exponential growth

In general, if you have enough information about a situation, you can write an exponential function to model growth in the situation. Let's start with a straightforward example:

Example 1: A social networking website is started by a group of 10 friends. They advertise their site before they launch, and membership grows fast: the membership doubles every day. At this rate, what will the membership be in a week? When will the membership reach 100,000?

Solution: To model this situation, let's look at how the membership changes each day:

Table 4.59:

Time (in days)	Membership
0	10
1	$2 \times 10 = 20$

Table 4.59: (continued)

Time (in days)	Membership
2	$2 \times 2 \times 10 = 40$
3	$2 \times 2 \times 2 \times 10 = 80$
4	$2 \times 2 \times 2 \times 2 \times 10 = 160$

Notice that the membership on day x is $10(2^x)$. Therefore we can model membership with the function $M(x) = 10(2^x)$. In seven days, the membership will be $M(7) = 10(2^7) = 1280$.

We can solve an exponential equation to find out when the membership will reach 100,000:

$$\begin{aligned} 10(2^x) &= 100,000 \\ 2^x &= 10,000 \\ log2^x &= log10,000 \\ xlog2 &= 4 \\ x &= \frac{4}{log2} \approx 13.3 \end{aligned}$$

At this rate, the membership will reach 100,000 in less than two weeks. This result may seem unreasonable. That's very fast growth!

So let's consider a slower rate of doubling. Let's say that the membership doubles every 7 days.

Table 4.60:

Time (in days)	Membership
0	10
7	$2 \times 10 = 20$
14	$2 \times 2 \times 10 = 40$
21	$2 \times 2 \times 2 \times 10 = 80$
28	$2 \times 2 \times 2 \times 2 \times 10 = 160$

We can no longer use the function $M(x) = 10(2^x)$. However, we *can* use this function to find another function to model this new situation. Looking at one data point will help. Consider for example the fact that $M(21) = 10(2^3)$. This is the case because 21 days results in 3 periods of doubling. In order for $x = 21$ to produce 2^3 in the equation, the exponent in the function must be $x/7$. So we have $M(x) = 10\left(2^{\frac{x}{7}}\right)$. Let's verify that this equation makes sense for the data in the table:

$$\begin{aligned} M(0) &= 10\left(2^{\frac{0}{7}}\right) = 10(1) = 10 \\ M(7) &= 10\left(2^{\frac{7}{7}}\right) = 10(2) = 20 \\ M(14) &= 10\left(2^{\frac{14}{7}}\right) = 10(2^2) = 10(4) = 40 \\ M(21) &= 10\left(2^{\frac{21}{7}}\right) = 10(2^3) = 10(8) = 80 \\ M(28) &= 10\left(2^{\frac{28}{7}}\right) = 10(2^4) = 10(16) = 160 \end{aligned}$$

Notice that each x value represents one more event of doubling, and in order for the function to have the correct power of 2, the exponent must be $(x/7)$.

With the new function $M(x) = 10\left(2^{\frac{x}{7}}\right)$, the membership doubles to 20 in one week, and reaches 100,000 in about 3 months:

$$10\left(2^{\frac{x}{7}}\right) = 100,000$$
$$2^{\frac{x}{7}} = 100,000$$
$$log2^{\frac{x}{7}} = log100,000$$
$$\frac{x}{7}log2 = 4$$
$$xlog2 = 28$$
$$x = \frac{28}{log2} \approx 93$$

The previous two examples of exponential growth have specifically been about doubling. We can also model a more general growth pattern with a more general growth model. While the graphing calculator produces a function of the form $y = a(b^x)$, population growth is often modeled with a function in which e is the base. Let's look at this kind of example:

The population of a town was 20,000 in 1990. Because of its proximity to technology companies, the population grew to 35,000 by the year 2000. If the growth continues at this rate, how long will it take for the population to reach 1 million?

The general form of the exponential growth model is much like the continuous compounding function you learned in the previous lesson. We can model exponential growth with a function of the form $P(\mathrm{t}) = P_0 e^{kt}$. The expression $P(t)$ represents the population after t years, the coefficient P_0 represents the initial population, and k is a growth constant that depends on the particular situation.

In the situation above, we know that $P_0 = 20{,}000$ and that $P(10) = 35,\ 000$. We can use this information to find the value of k:

$$P(t) = P_0 e^{kt}$$
$$P(10) = 35000 = 20000e^{k \cdot 10}$$
$$\frac{35,000}{20000} = e^{10k}$$
$$1.75 = e^{10k}$$
$$ln1.75 = lne^{10k}$$
$$ln1.75 = 10klne$$
$$ln1.75 = 10k(1)$$
$$ln1.75 = 10k$$
$$k = \frac{ln1.75}{10} \approx 0.056$$

Therefore we can model the population growth with the function $P(t) = 20000e^{\frac{ln1.75}{10}t}$. We can determine when the population will reach 1,000,000 by solving an equation, or using a graph.

Here is a solution using an equation:

$$1000000 = 20000e^{\frac{ln1.75}{10}t}$$
$$50 = e^{\frac{ln1.75}{10}t}$$
$$ln50 = ln\left(e^{\frac{ln1.75}{10}t}\right)$$
$$ln50 = \frac{ln1.75}{10}t(lne)$$
$$ln50 = \frac{ln1.75}{10}t(1)$$
$$10ln50 = ln1.75t$$
$$t = \frac{10ln50}{ln1.75} \approx 70$$

At this rate, it would take about 70 years for the population to reach 1 million. Like the initial doubling example, the growth rate may seem very fast. In reality, a population that grows exponentially may not sustain its growth rate over time. Next we will look at a different kind of function that can be used to model growth of this kind.

Logistic models

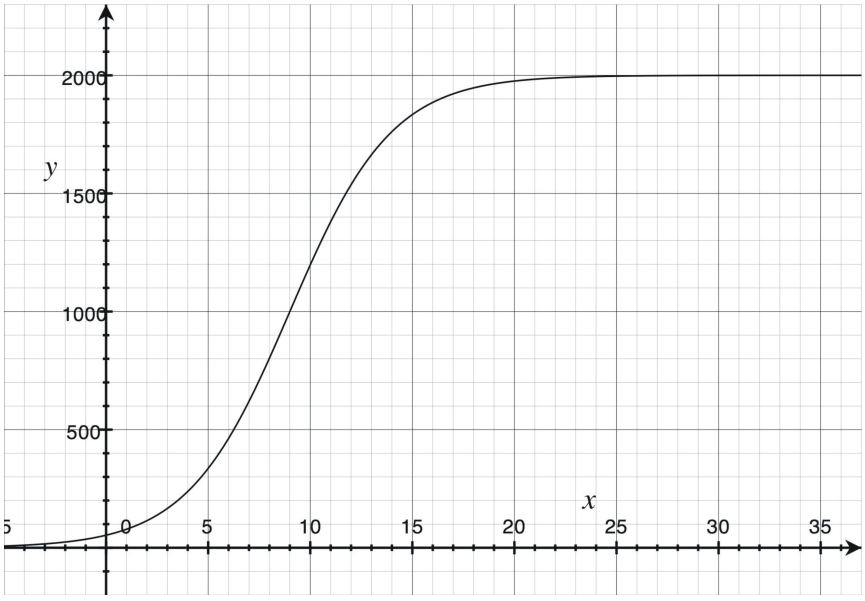
Given that resources are limited, a population may slow down in its growth over time. Consider the last example, the town whose population exploded in the 1990s. If there are no more houses to be bought, or tracts of land to be developed, the population will not continue to grow exponentially. The table below shows the population of this town slowing down, though still growing:

Table 4.61:

t (1990=0)	Population
0	20,000
10	35,000
15	38,000
20	40,000

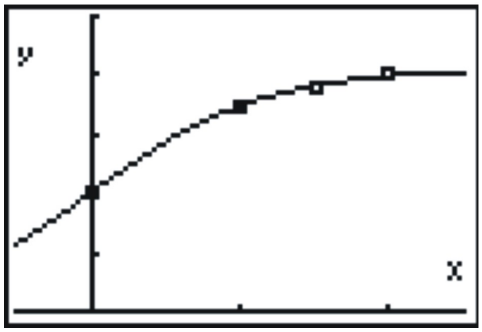
As the population growth slows down, the population may approach what is called a **carrying capacity,** or an upper bound of the population. We can model this kind of growth using a logistic function, which is a function of the form $f(x) = \frac{c}{1+a(e^{-bx})}$.

The graph below shows an example of a logistic function. This kind of graph is often called an "s curve" because of its shape.



Notice that the graph shows slow growth, then fast growth, and then slow growth again, as the population or quantity in question approaches the carrying capacity. Logistics functions are used to model population growth, as well as other situations, such as the amount of medicine in a person's system

Given the population data above, we can use a graphing calculator to find a logistic function to model this situation. The details of this process are explained in the Technology Note in Lesson 3.5. As shown there, enter the data into L_1 and L_2. Then run a logistic regression. (Press <TI font_STAT>, scroll right to CALC, and scroll down to B. Logistic.) An approximation of the logistic model for this data is: $f(x) = \frac{41042.38}{1+1.050e^{-.178x}}$. A graph of this function and the data is shown here.



Notice that the graph has a horizontal asymptote around 40,000. Looking at the equation, you should notice that the numerator is about 41,042. This value is in fact the horizontal asymptote, which represents the carrying capacity. We can understand why this is the carrying capacity if we consider the limit of the function as x approaches infinity. As x gets larger and larger, $e^{-.178x}$ will get smaller and smaller. So 1.05 $e^{-.178x}$ will get smaller. This means that the denominator of the function will get closer and closer to 1:

$\lim_{x\to\infty}(1 + 1.050e^{-.178x}) = 1.$

Therefore the limit of the function is (approximately) (41042/1) = 41042. This means that given the current growth, the model predicts that the population will not go beyond 41,042. This kind of growth is seen in population, as well as other situations in which some quantity grows very fast and then slows down, or when a quantity steeply decreases, and then levels off. You will see work with more examples of logistic functions in the review questions.

Exponential decay

Just as a quantity can grow, or increase exponentially, we can model a decreasing quantity with an exponential function. This kind of situation is referred to as exponential decay. Perhaps the most common example of exponential decay is that of radioactive decay, which refers to the transformation of an atom of one type into an atom of a different type, when the nucleus of the atom loses energy. The rate of radioactive decay is usually measured in terms of "half-life," or the time it takes for half of the atoms in a sample to decay. For example Carbon-14 is a radioactive isotope that is used in "carbon dating," a method of determining the age of organic materials. The half-life of Carbon-14 is 5730 years. This means that if we have a sample of Carbon-14, it will take 5730 years for half of the sample to decay. Then it will take another 5730 years for half of the remaining sample to decay, and so on.

We can model decay using the same form of equation we use to model growth, except that the exponent in the equation is negative: $A(t) = A_0\ e^{-kt}$. For example, say we have a sample of Carbon-14. How much time will pass before 75% of the original sample remains?

We can use the half-life of 5730 years to determine the value of k:

Table 4.62:

$A(t) = A_0e^{-kt}$	
$\frac{1}{2} = 1e^{-k\cdot 5730}$	We do not know the value of A_0, *so we use "1" as 100%. (1/2) of the sample remains when* t *= 5730 years*
$ln\frac{1}{2} = lne^{-k\cdot 5730}$	Take the ln of both sides

Table 4.62: (continued)

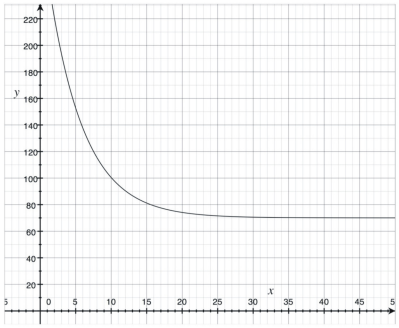
$A(t) = A_0 e^{-kt}$	
$ln\frac{1}{2} = -5730klne$	Use the power property of logs
$ln\frac{1}{2} = -5730k$	$ln(e) = 1$
$-ln2 = -5730k$	$ln(1/2) = ln(2^{-1}) = -ln2$
$ln2 = 5730k$	
$k = \frac{ln2}{5730}$	Isolate k

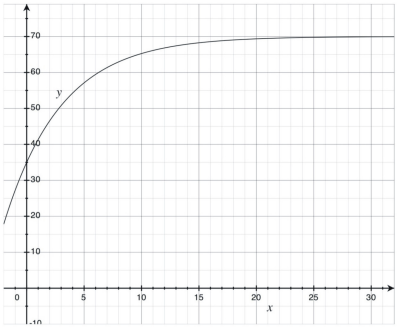
Now we can determine when the amount of Carbon-14 remaining is 75% of the original:

$$0.75 = 1e^{\frac{-ln2}{5730}t}$$
$$0.75 = 1e^{\frac{-ln2}{5730}t}$$
$$ln(0.75) = lne^{\frac{-ln2}{5730}t}$$
$$ln(0.75) = \frac{-ln2}{5730}t$$
$$t = \frac{5730ln(0.75)}{-ln2} \approx 2378$$

Therefore it would take about 2,378 years for 75% of the original sample to be remaining. In practice, scientists can approximate the age of an artifact using a process that relies on their knowledge of the half-life of Carbon-14, as well as the ratio of Carbon-14 to Carbon-12 (the most abundant, stable form of carbon) in an object. While the concept of half-life often is used in the context of radioactive decay, it is also used in other situations. In the review questions, you will see another common example, that of medicine in a person's system.

Related to exponential decay is Newton's Law of Cooling. The Law of Cooling allows us to determine the temperature of a cooling (or warming) object, based on the temperature of the surroundings and the time since the object entered the surroundings. The general form of the cooling function is $T(x) = T_5 + (T_0 - T_5)\ e^{-kx}$, where T_5, is the surrounding temperature, T_0 is the initial temperature, and x represents the time since the object began cooling or warming.





The first graph shows a situation in which an object is cooling. The graph has a horizontal asymptote at $y = 70$. This tells us that the object is cooling to 70°F. The second graph has a horizontal asymptote at $y = 70$ as well, but in this situation, the object is warming up to 70°F.

We can use the general form of the function to answer questions about cooling (or warming) situations. Consider the following example: you are baking a casserole in a dish, and the oven is set to 325°F. You take the pan out of the oven and put it on a cooling rack in your kitchen which is 70°F, and after 10 minutes the pan has cooled to 300°F. How long will it take for the pan to cool to 200^0F?

We can use the general form of the equation and the information given in the problem to find the value of k:

$$T(x) = T_s + (T_0 - T_s)e^{-kx}$$
$$T(x) = 70 + (325 - 70)e^{-kx}$$
$$T(x) = 70 + (255)e^{-kx}$$
$$T(10) = 70 + 255e^{-10k} = 300$$
$$255e^{-10k} = 230$$
$$e^{-10k} = \frac{230}{255}$$
$$lne^{-10k} = ln\left(\frac{230}{255}\right)$$
$$-10k = ln\left(\frac{230}{255}\right)$$
$$k = \frac{ln\left(\frac{230}{255}\right)}{-10} \approx 0.0103$$

Now we can determine the amount of time it takes for the pan to cool to 200 degrees:

$$T(x) = 70 + (255)e^{-.0103}x$$
$$T(x) = 70 + (255)e^{-.0103}x$$
$$200 = 70 + (255)e^{-.0103}x$$
$$130 = (255)e^{-.0103}x$$
$$\frac{130}{255} = e^{-.0103}x$$
$$ln\left(\frac{130}{255}\right) = lne^{-.0103}x$$
$$ln\left(\frac{130}{255}\right) = -.0103x$$
$$x = \frac{ln\left(\frac{130}{255}\right)}{-.0103} \approx 65$$

Therefore, in the given surroundings, it would take about an hour for the pan to cool to 200 degrees.

Lesson Summary

In this lesson we have developed exponential and logistic models to represent different phenomena. We have considered exponential growth, logistic growth, and exponential decay. After reading the examples in this lesson, you should be able to write a function to represent a given situation, to evaluate the function for a given value of x, and to solve exponential equations in order to find values of x, given values of the function. For example, in a situation of exponential population growth as a function of time, you should be able to determine the population at a particular time, and to determine the time it takes for the population to reach a given amount. You should be able to solve these kinds of problems by solving exponential equations, and by using graphing utilities, as we have done throughout the chapter.

Points to Consider

1. How can we use the same equation for exponential growth and decay?
2. What are the restrictions on domain and range for the examples in this lesson?
3. How can we use different equations to model the same situations?

Review Questions

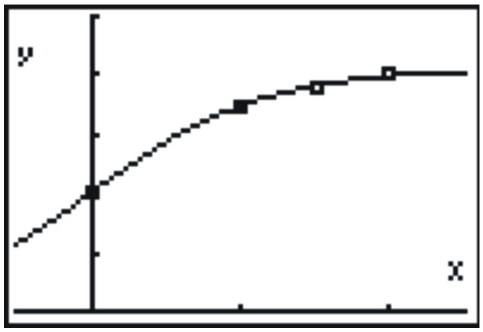
1. The population of a town was 50,000 in 1980, and it grew to 70,000 by 1995. a. Write an exponential function to model the growth of the population.

b. Use the function to estimate the population in 2010.
c. What if the population growth was linear? Write a linear equation to model the population growth, and use it to estimate the population in 2010.

2. A telecommunications company began providing wireless service in 1994, and during that year the company had 1000 subscribers. By 2004, the company had 12,000 subscribers. a. Write an exponential function to model the situation
b. Use the model to determine how long it will take for the company to reach 50,000 subscribers.
3. The population of a particular strain of bacteria triples every 8 hours. a. Write a general exponential function to model the bacteria growth.
b. Use the model to determine how long it will take for a sample of bacteria to be 100 times its original population.
c. Use a graph to verify your solution to part b.
4. The half-life of acetaminophen is about 2 hours. a. If you take 650 mg of acetaminophen, how much will be left in your system after 7 hours?
b. How long before there is less than 25 mg in your system?
5. The population of a city was 200,000 in 1991, and it decreased to 170,000 by 2001. a. Write an exponential function to model the decreasing population, and use the model to predict the population in 2008.
b. Under what circumstances might the function cease to model the situation after a certain point in time?
6. Consider the following situation: you buy a large box of pens for the start of the school year, and after six weeks, (1/3) of the pens remain. After another six weeks, (1/3) of the remaining pens were remaining. If you continue this pattern, when will you only have 5% of the pens left?
7. Use Newton's law of cooling to answer the question: you pour hot water into a mug to make tea. The temperature of the water is about 200 degrees. The surrounding temperature is about 75 degrees. You let the water cool for 5 minutes, and the temperature decreases to 160 degrees. What will the temperature be after 15 minutes?
8. The spread of a particular virus can be modeled with the logistic function $f(x) = \frac{2000}{1+600e^{-.75x}}$, where x is the number of days the virus has been spreading, and f(x) represents the number of people who have the virus.a. How many people will be affected after 7 days?
b. How many days will it take for the spread to be within one person of carrying capacity?
9. Consider again the situation in problem #2: A telecommunications company began providing wireless service in 1994, and during that year the company had 1000 subscribers. By 2004, the company had 12,000 subscribers. If the company has 15,000 subscribers in 2005, and 16,000 in 2007, what type of model do you think should be used to model the situation? Use a graphing calculator to find a regression equation, and use the equation to predict the number of subscribers in 2010.
10. Compare exponential and logistic functions as tools for modeling growth. What do they have in common, and how do they differ?

Review Answers

1. a. $A(T) = 50,000e^{\frac{ln(\frac{7}{5})}{15}t}$b. $98,000$
c. $f(t) = \frac{4000}{3}t + 50000$. The population would be 90,000, which is different by about 9%.
2. a. $S(t) = 1000e^{\frac{ln(12)}{10}t}$b. $t = \frac{10ln50}{ln12} \approx 15.74$
3. a. $A(t) = A_0\ (3^{t/8})$b. $t = \frac{16}{log3} \approx 33.53$
c. The graph below shows $y = 100$ and $y = 3^{x/8}$, which intersect at approximately $x = 33.53$



4. a. About 57.45 mg b. About 9.4 hours
5. a. $P(t) = 200000e^{\frac{-ln.85}{-10}t}, P(17) \approx 151720$b. If the economy or other factors change, the population might begin to increase, or the rate of decrease could change as well.
6. $t = \frac{6log(0.05)}{log(\frac{1}{3})} \approx 16$ weeks
7. About 114 degrees.
8. a. About 482 people. b. After 19 days, over 1999 people have the virus.
9. The graph indicates a logistic model. $f(x) \approx \frac{18872}{1+21.45e^{-377x}}$ gives 17952 subscribers in 2010.
10. Both types of functions model fast increase in growth, but the logistic model shows the growth slowing down after some point, with some upper bound on the quantity in question. (Many people argue that logistic growth is more realistic.)

Vocabulary

Carrying capacity The supportable population of an organism, given the food, habitat, water and other necessities available within an ecosystem is known as the ecosystem's carrying capacity for that organism.

Radioactive decay Radioactive decay is the process in which an unstable atomic nucleus loses energy by emitting radiation in the form of particles or electromagnetic waves. This decay, or loss of energy, results in an atom of one type transforming to an atom of a different type. For example, Carbon-14 transforms into Nitrogen-14

Half-life The amount of time it takes for half of a given amount of a substance to decay. The half-life remains the same, no matter how much of the substance there is.

Isotope Isotopes are any of two or more forms of a chemical element, having the same number of protons in the nucleus, or the same atomic number, but having different numbers of neutrons in the nucleus, or different atomic weights.

4.8 Applications

Learning objectives

- Work with the decibel system for measuring loudness of sound.
- Work with the Richter scale, which measures the magnitude of earthquakes.
- Work with pH values and concentrations of hydrogen ions.

Introduction

Because logarithms are related to exponential relationships, logarithms are useful for measuring phenomena that involve very large numbers or very small numbers. In this lesson you will learn about three situations in which a quantity is measured using logarithms. In each situation, a logarithm is used to simplify measurements of either very small numbers or very large numbers. We begin with measuring the intensity of sound.

Intensity of sound

Sound intensity is measured using a logarithmic scale. The intensity of a sound wave is measured in Watts per square meter, or W/m^2. Our hearing threshold (or the minimum intensity we can hear at a frequency of 1000 Hz), is $2.5 \times 10^{12}\ W/m^2$. The intensity of sound is often measured using the decibel (dB) system. We can think of this system as a function. The input of the function is the intensity of the sound, and the output is some number of decibels. The decibel is a dimensionless unit; however, because decibels are used in common and scientific discussions of sound, the values of the scale have become familiar to people.

We can calculate the decibel measure as follows:

$$\text{Intensity level } (dB) = 10log\left[\frac{\text{intensity of sound in } W/m^2}{.937\times10^{-12}W/m^2}\right]$$

An intensity of $.937 \times 10^{-12}\ W/m^2$ corresponds to 0 decibels:

$$10log\left[\frac{.937\times10^{-12}W/m^2}{.937\times10^{-12}W/m^2}\right] = 10log1 = 10(0) = 0.$$

Note: The sound equivalent to 0 decibels is approximately the lowest sound that humans can hear. If the intensity is ten times as large, the decibel level is 10:

$$10log\left[\frac{.937\times10^{-11}W/m^2}{.937\times10^{-12}W/m^2}\right] = 10log10 = 10(1) = 10$$

If the intensity is 100 times as large, the decibel level is 20, and if the intensity is 1000 times as large, the decibel level is 30. (The scale is created this way in order to correspond to human hearing. We tend to underestimate intensity.) The threshold for pain caused by sound is 1 W/m^2. This intensity corresponds to about 120 decibels:

$$10log\left[\frac{1W/m^2}{.937\times10^{-12}}W/m^2\right] \approx 10:12 = 120$$

Many common phenomena are louder than this. For example, a jet can reach about 140 decibels, and concert can reach about 150 decibels.

(Source: Ohanian, H.C. (1989) Physics. New York: W.W. Norton & Company.)

For ease of calculation, the equation is often simplified: .937 is rounded to 1:

Table 4.63:

Intensity level (dB)	$= 10log\left[\frac{\text{intensity of sound in } W/m^2}{1\times10^{-12}W/m^2}\right]$
	$= 10log\left[\frac{\text{intensity of sound in } W/m^2}{10^{-12}W/m^2}\right]$

In the example below we will use this simplified equation to answer a question about decibels. (In the review exercises, you can also use this simplified equation).

Example 1: Verify that a sound of intensity 100 times that of a sound of 0 dB corresponds to 20 dB.

Solution: $dB = 10log\left(\frac{100\times10^{-12}}{10^{-12}}\right) = 10log(100) = 10(2) = 20$.

Intensity and magnitude of earthquakes

An earthquake occurs when energy is released from within the earth, often caused by movement along fault lines. An earthquake can be measured in terms of its intensity, or its magnitude. Intensity refers to the effect of the earthquake, which depends on location with respect to the epicenter of the quake. Intensity and magnitude are not the same thing.

As mentioned in lesson 3, the magnitude of an earthquake is measured using logarithms. In 1935, scientist Charles Richter developed this scale in order to compare the "size" of earthquakes. You can think of Richter scale as a function in which the input is the amplitude of a seismic wave, as measured by a seismograph, and the output is a magnitude. However, there is more than one way to calculate the magnitude of an earthquake because earthquakes produce two different kinds of waves that can be measured for amplitude. The calculations are further complicated by the need for a correction factor, which is a function of the distance between the epicenter and the location of the seismograph.

Given these complexities, seismologists may use different formulas, depending on the conditions of a specific earthquake. This is done so that the measurement of the magnitude of a specific earthquake is consistent with Richter's original definition. (Source: http://earthquake.usgs.gov/learning/topics/richter.php)

Even without a specific formula, we can use the Richter scale to compare the size of earthquakes. For example, the 1906 San Francisco earthquake had a magnitude of about 7.7. The 1989 Loma Prieta earthquake had a magnitude of about 6.9. (The epicenter of the quake was near Loma Prieta peak in the Santa Cruz mountains, south of San Francisco.) Because the Richter scale is logarithmic, this means that the 1906 quake was six times as strong as the 1989 quake:

$$\frac{10^{7.7}}{10^{6.9}} = 10^{7.7-6.9} = 10^{.8} \approx 6.3$$

This kind of calculation explains why magnitudes are reported using a whole number and a decimal. In fact, a decimal difference makes a big difference in the size of the earthquake, as shown below and in the review exercises

Example 2: An earthquake has a magnitude of 3.5. A second earthquake is 100 times as strong. What is the magnitude of the second earthquake?

Solution: The second earthquake is 100 times as strong as the earthquake of magnitude 3.5. This means that if the magnitude of the second earthquake is x, then:

$$\frac{10^{x}}{10^{3.5}} = 100$$

$10^{x-3.5} = 100 = 10^2$
$x - 3.5 = 2$
$x = 5.5$

So the magnitude of the second earthquake is 5.5.

The pH scale

If you have studied chemistry, you may have learned about acids and bases. An acid is a substance that produces hydrogen ions when added to water. A hydrogen ion is a positively charged atom of hydrogen, written as H^+. A base is a substance that produces hydroxide ions (OH^-) when added to water. Acids and bases play important roles in everyday life, including within the human body. For example, our stomachs produce acids in order to breakdown foods. However, for people who suffer from gastric reflux, acids travel up to and can damage the esophagus. Substances that are bases are often used in cleaners, but a strong base is dangerous: it can burn your skin.

To measure the concentration of an acid or a base in a substance, we use the pH scale, which was invented in the early 1900's by a Danish scientist named Soren Sorenson. The pH of a substance depends on the concentration of H^+, which is written with the symbol $[H^+]$.

$pH = - \log [H^+]$

(Note: concentration is usually measured in moles per liter. A mole is 6.02×10^{23} units. Here, it would be 6.02×10^{23} hydrogen ions.)

For example, the concentration of H^+ in stomach acid is about 1×10^{-1}. So the pH of stomach acid is -log $(10^{-1}) = -(-1) = 1$. The pH scale ranges from 0 to 14. A substance with a low pH is an acid. A substance with a high pH is a base. A substance with a pH in the middle of the scale is considered to be neutral.

Example 3: The pH of ammonia is 11. What is the concentration of H^+?

Solution: $pH = - \log [H^+]$. If we substitute 11 for pH we can solve for H^+:

$11 = -log[H^+]$
$-11 = log[H^+]$
$10^{-11} = 10^{log[H^+]}$
$10^{-11} = H^+$

Lesson Summary

In this lesson we have looked at three examples of logarithmic scales. In the case of the decibel system, using a logarithm has produced a simple way of categorizing the intensity of sound. The Richter scale allows us to compare earthquakes. And, the pH scale allows us to categorize acids and bases. In each case, a logarithm helps us work with large or small numbers, in order to more easily understand the quantities involved in certain real world phenomena.

Points to Consider

1. How are the decibel system and the Richter scale the same, and how are they different?
2. What other phenomena might be modeled using a logarithmic scale?

Review Questions

1. Verify that a sound of intensity 1000 times that of a sound of 0 dB corresponds to 30 dB.
2. Calculate the decibel level of a sound with intensity 10^{-8} W/m^2.
3. Calculate the intensity of a sound if the decibel level is 25.
4. The 2004 Indian Ocean earthquake was recorded to have a magnitude of about 9.5. In 1960, an earthquake in Chile was recorded to have a magnitude of 9.1. How much stronger was the 2004 Indian Ocean quake?
5. Two earthquakes of the same magnitude do not necessarily cause the same amount of destruction. How is that possible?
6. The concentration of H^+ in pure water is 1×10^{-7}. What is the pH?
7. The pH of normal human blood is 7.4. What is the concentration of H^+

Review Answers

1. $dB = 10log\left(\frac{100 \times 10^{-12}}{10^{-12}}\right) = 10log(100) = 10(3) = 30$
2. $dB = 10log\left(\frac{10^{-8}}{10^{-12}}\right) = 10log(10000) = 10(4) = 40$
3. $10^{-9.5}$ or 3.16×10^{-10}
4. $10^{0.4} \approx 2.5$
5. According to the USGS, the damage depends on the strength of shaking, the length of shaking, the type of soil in the area, and the types of buildings. Many buildings in the San Francisco Bay Area are undergoing "seismic retrofitting," in anticipation of "the big one."
6. The pH is 7.
7. $10^{-7.4} \approx 3.98 \times 10^{-8}$

Vocabulary

Acid An acid is a substance that produces hydrogen ions when added to water.

Amplitude The amplitude of a wave is the distance from its highest (or lowest) point to its center.

Base A base is a substance that produces hydroxide ions (OH^-) when added to water

Decibel A decibel is a unitless measure of the intensity of sound.

Mole 6.02×10^{23} units of a substance.

Seismograph A seismograph is a device used to measure the amplitude of earthquakes.

Chapter 5

Polar Equations and Complex Numbers

5.1 Polar Coordinates

Learning Objectives

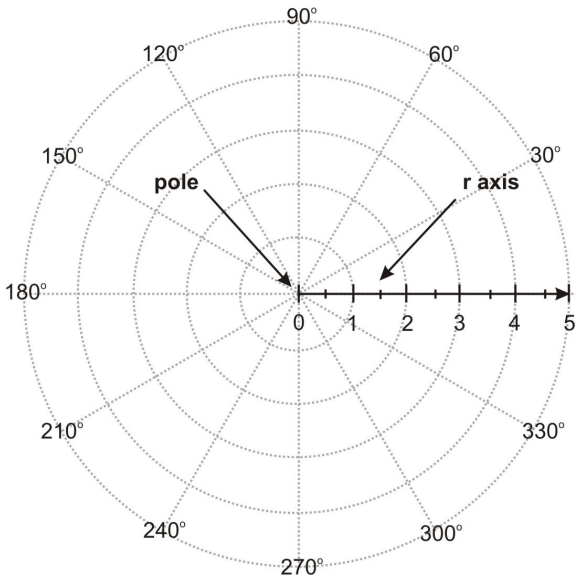
- Identify major parts of the polar coordinate system: the pole, polar axis, , r
- Plot points in the polar coordinate system
- Use a graphing utility to plot common polar functions

Pole, Polar Axis, , r

Symmetry can be explored by using trigonometric functions and a coordinate system called the Polar Coordinate System.

The Polar Coordinate System consists of a ray known as the Polar Axis and the endpoint of the ray, called the Pole.

The coordinate $(r, \ )$, where r is a distance from the Pole and is the angle a line segment from the pole to the point makes with the Polar Axis. The point $(r, \ )$ is known as the Polar Coordinate.

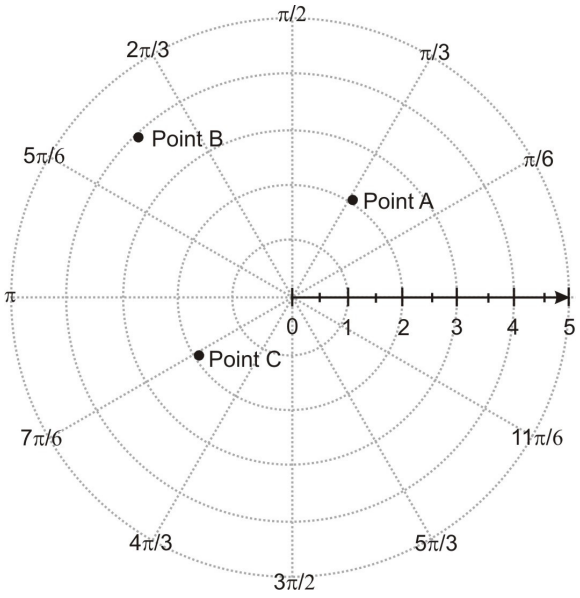


Plot (, r)

To plot a point, first go along the r-axis by r units. Then, rotate counterclockwise by the angle . Be careful to use the correct units for the angle measure (either radians or degrees).

Examples of polar coordinates are point A $\left(2, \frac{\pi}{3}\right)$, point B $(4, 135^{\circ})$, and point C $\left(-2, \frac{\pi}{6}\right)$.

Below is the pole, polar axis and the points A, B and C.



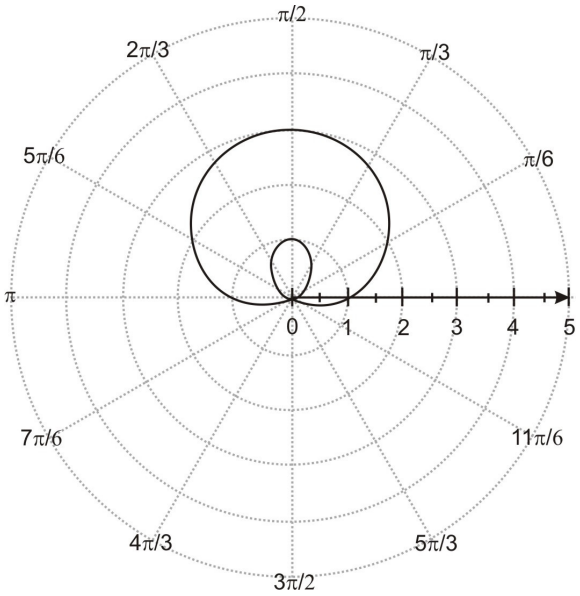
Usually polar plots are done with radians (especially if they include trigonometric functions), but sometimes degrees are used.

If r is negative you can plot a point $(r, \ )$ by first plotting $(|\ r\ |, \ )$, and then rotating by $= 180^{\circ}$.

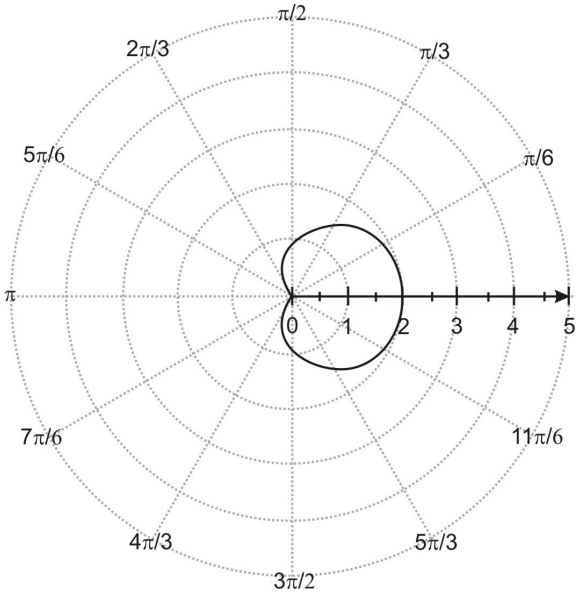
Sinusoids of One Revolution (e.g., limaçons, cardioids)

As with the rectangular coordinate system, equations can be graphed in the polar coordinate system. We will go into more detail on graphing in a future section. For now, these graphs are examples of what shapes are possible to graph as polar functions.

Below is the graph of a polar equation $r = 1 + 2 \sin$ (limaçon)



Graph: $r = 1 + \cos$ (cardioid)



Applications, Technological Tools

Polar Equations can be graphed using a graphing calculator: With the graphing calculator- go to **MODE**. There select **RADIAN** for the angle measure and **POL** (for Polar) on the **FUNC** (function)line. When Y = is pressed, note that the equation has changed from y = to r = . There input the polar equation. After pressing graph, if you can't see the full graph, adjust x- and y- max/min, etc in **WINDOW.**

Lesson Summary

Until now, most, if not all, graphing has been done using the rectangular coordinate system. There are other coordinate systems, such as polar coordinates, with which graphing can be accomplished.

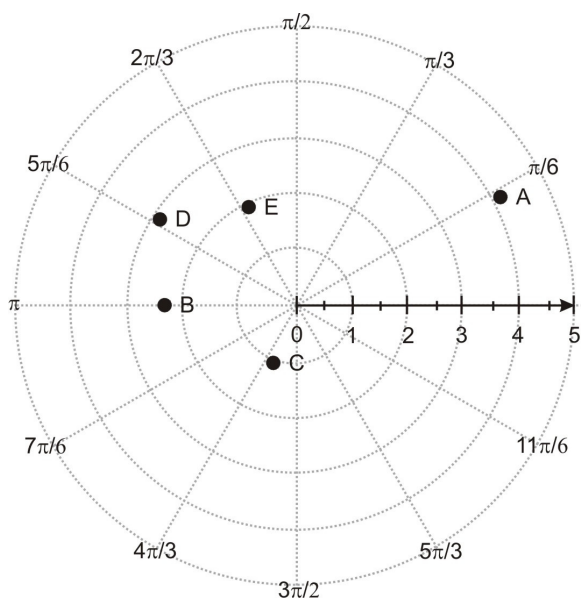
Points to Consider

Some equations are easier to graph in one coordinate system as opposed to the other coordinate system. Knowing how to graph in each coordinate system allow the freedom to be able to graph any equation.

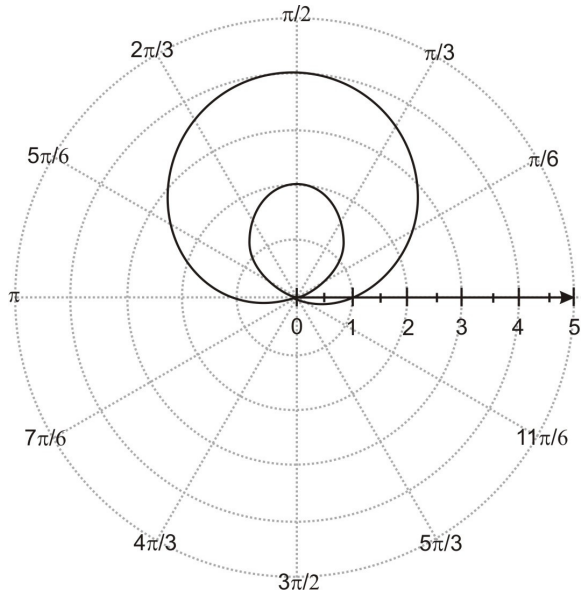
Review Questions

1. Plot the following points: a. (4, 30°)
   b. (2.5,)
   c. $\left(-1, \frac{\pi}{3}\right)$
   d. $\left(3, \frac{5\pi}{6}\right)$
   e. (-2, 300°)
2. Use a graphing calculator or plotting program to plot the following equations: a. 1 + 3 sin
   b. 1 + 2 cos

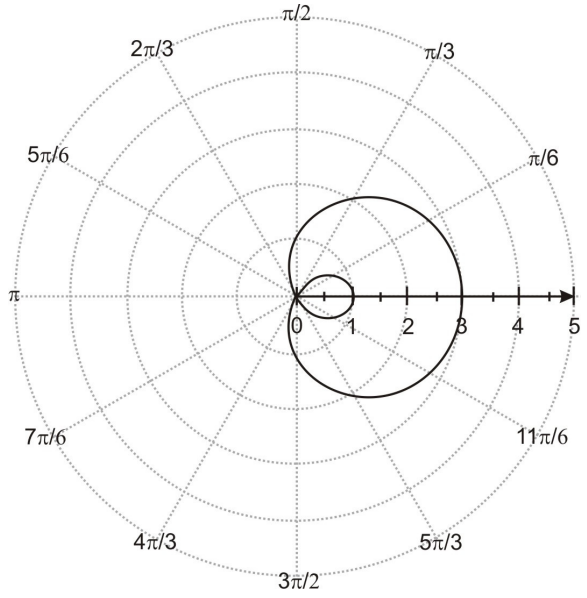
Review Answers

1. 

2. a.



b.



5.2 Polar-Cartesian Transformations

Learning Objectives

- Convert polar points to rectangular form
- Convert rectangular points to polar form
- Convert some graphs from rectangular to polar

Introduction

You have learned that the polar coordinate system offers a way to plot points in a plane. Also, you have extensive experience using the rectangular (x - y) coordinate system. Sometimes it is helpful to transform a point or a whole graph from one coordinate system to the other.

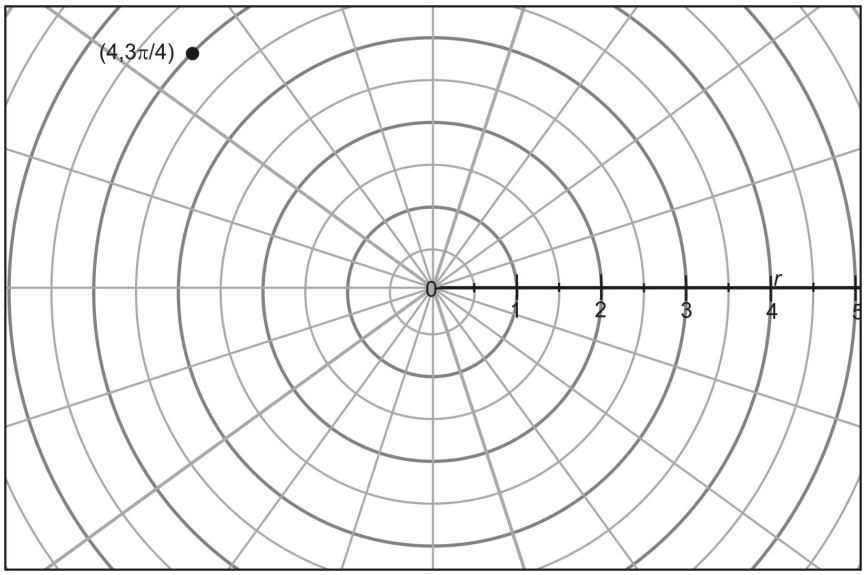
In this section you will learn how to convert polar points to rectangular, and vice versa. You will also see some surprising ways to write conic sections in polar form.

Polar to Rectangular

Sometimes a problem will be given as a coordinate in polar form but rectangular form may be needed.

To transform the polar point $\left(4, \frac{3\pi}{4}\right)$ into rectangular coordinates: first identify (r,)

r = 4 and $\theta = \frac{3\pi}{4}$.



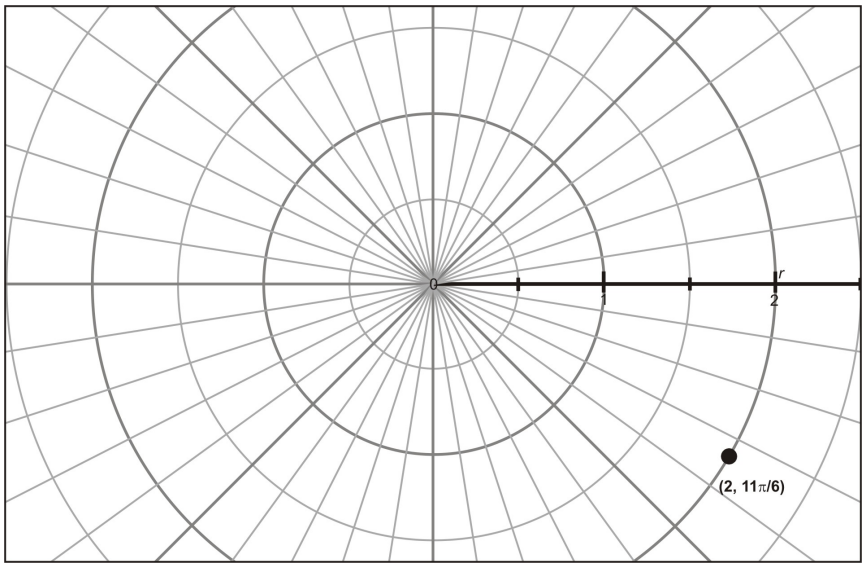
Second, draw the point on polar axis. Drop a line from the point to the polar axis. The distance from the pole to where the line intersects the polar axis is the x value, and the length of the line segment from the point to the polar axis is the y value:

$x = r \cos$ and $y = r \sin$

$x = 4 \cos \frac{3\pi}{4}$ and $y = 4 \sin \frac{3\pi}{4}$ or $x = -2\sqrt{2}$ $y = 2\sqrt{2}$

$\left(4, \frac{3\pi}{4}\right)$ in polar coordinates is equivalent to $(-2\sqrt{2},\ 2\sqrt{2})$ in rectangular coordinates

Example: Transform the polar coordinates $\left(2, \frac{11\pi}{6}\right)$ to rectangular form



$r = 2$ and $\theta = \frac{11\pi}{6}$

$x = r \cos \theta$ and $y = r \sin \theta$

$x = 2 \cos \frac{11\pi}{6}$ and $y = 2 \sin \frac{11\pi}{6}$ or $x = 3\sqrt{2} \quad y = -1$

$\left(2, \frac{11\pi}{6}\right)$ is equivalent to $(3\sqrt{2}, -1)$ or in decimal form, approximately (4.342, -1)

Rectangular to Polar

Going from rectangular coordinates to polar coordinates is also possible, but it takes a bit more work. Suppose we want t find the polar coordinates of the rectangular point (2, 2). To begin doing this operation, the distance that the point (2, 2) is from the origin (the radius, r) can be found by

$r = \sqrt{x^2 + y^2}$

$r = \sqrt{2^2 + 2^2}$

$r = \sqrt{8} = 2\sqrt{2}$

The angle that the line segment between the point and the origin can be found by

$\tan \theta = \frac{y}{x}$

$\tan \theta = \frac{2}{2}$

$\tan \theta = 1$

$\theta = \tan^{-1} 1$

$\theta = \frac{\pi}{4}$

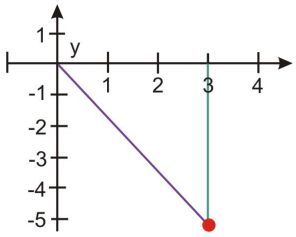
Since this point is in the first quadrant (both the x and y coordinate are positive) the angle must be 45^o or $\frac{\pi}{4}$ radians. It is also possible that when $\tan \quad = 1$ the angle can be in the third quadrant, or $\frac{5\pi}{4}$ radians. But this angle will not satisfy the conditions of the problem, since a third quadrant angle must have both x and y negative.

Note: when you use using $\tan \theta = \frac{y}{x}$ to find the measure of you should consider, at first, the quotient $\tan \theta = \left|\frac{y}{x}\right|$ and find the first quadrant angle that satisfies this condition. This angle will be called the reference angle, denoted $_{ref}$. Find the actual angle by analyzing which quadrant the angle must be given the signs of x and y.

Example: Find the Polar coordinates for $(3, -3\sqrt{3})$

$x = 3$ and $y = -3\sqrt{3}$

Now, draw a right triangle in standard form. Find the distance the point is from the origin and the angle the line segment that represents this distance makes with the +x axis:



$r = \sqrt{3^2 + (-3\sqrt{3})^2}$

$= \sqrt{9 + 27}$

$= \sqrt{36}$

$= 6$

And for the angle,

$\tan \theta_{ref} = \left|\frac{(-3\sqrt{3})}{3}\right|$

$\tan \theta_{ref} = \sqrt{3}$

$\theta_{ref} = \tan^{-1} \sqrt{3}$

$\theta_{ref} = \frac{\pi}{3}$

So, $\theta_{ref} = \frac{\pi}{3}$ and we can look at the signs of x and y,– (+, -) – to see that $\theta = \frac{5\pi}{3}$ since it is a 4^{th} quadrant angle.

The rectangular point $(3, -3\sqrt{3})$ is equivalent to the polar point $\left(6, \frac{5\pi}{3}\right)$.

Recall that when solving for , we used

$\tan \theta = \left|\frac{(-3\sqrt{3})}{3}\right|$ or $\tan \theta = \sqrt{3}$

We found

$\theta = \frac{5\pi}{3}$. BUT, could also be $\theta = \frac{2\pi}{3}$. You must examine the signs of each coordinate to see that the angle must be in the fourth quadrant in rectangular units or between $\frac{3\pi}{2}$ and 2 in polar units. Of the two possible angles for , only $\frac{5\pi}{3}$ is valid. Note that when you use $\tan^{-1}$ on a calculator you will always get an answer in the range $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$.

Conic Section Transformations

Circle:

$x^2 + y^2 = k^2$ is the equation of a circle with a radius of k in rectangular coordinates. You can transform this equation to polar form by substituting the polar values for x, y. Recall $x = r \cos$ and $y = r \sin$

$(r \cos)^2 + (r \sin)^2 = k^2$

or by squaring the terms:

$r^2 \cos^2 + r^2 \sin^2 = k^2$

or by factoring the r^2 from both terms on the left

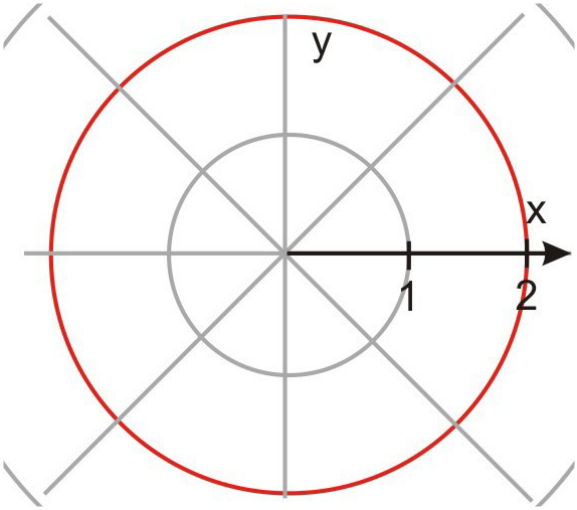
r^2 (cos^2 $\quad$ + sin^2 $\quad$) = k^2

and recalling the identity cos^2 $\quad$ + sin^2 $\quad$ = 1

$r^2 = k^2$

$r = \pm k$ is an equation for a circle in polar units

When r is equal to a constant, the polar graph is a circle. This was confirmed previously when introduced to polar graphing.



In the graph above of $r = 2$, at every angle the distance to the pole is 2 units. That is the meaning of $r = 2$. (Note that circles in the polar coordinate system are analogous to horizontal lines in the rectangular coordinate system).

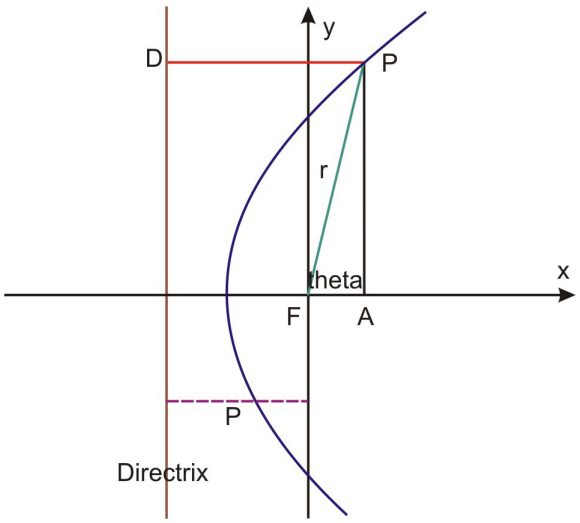
Optional: Parabola, Hyperbola and Ellipse in Polar form

We could convert the traditional equations for parabolas, hyperbolas, and ellipses into polar form using the same conversions

$x = r$ cos $\quad$ and $y = r$ sin

But, there is a much more elegant way to graph parabola, hyperbolas and ellipses in polar form. Below we introduce this method.

First, for these conic sections, recalling the definition of eccentricity will be useful.



$\frac{PF}{PD}$ = e (eccentricity).

e is the ratio of the distance between a point on the conic and its focus (PF) and the distance between the point on the conic and the conic's directrix (PD)

For convenience the focus is placed at the pole (0, 0)

The distance between the point on the conic and conic's focus (on the pole) is equal to r, and as seen in the diagram, the distance between the point on the conic and the conic's directrix, PD is equal to p + FA.

$e = \frac{r}{p+FA}$

Also in the diagram, note that $\frac{FA}{r} = \cos\theta$

$e = \frac{r}{p+r\cos\theta}$

solving for r:

$ep + er\cos\ = r$

or

$r - er\cos\ = ep$

factoring for r:

$r(1 - e\cos\ ) = ep$

and finally we get Equation A:

$r = \frac{ep}{1-e\cos\theta}$

In this equation the denominator contains 1 - e cos term

Examine the graph above: the conic opens to the right. This equation is valid for any conic that is opening to the right or if the major axis is the x-axis.

The following shows the other forms of conics in polar form:

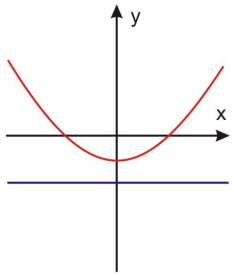
$r = \frac{ep}{1-e\sin\theta}$

opens to up, and

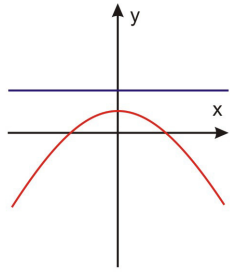
$r = \frac{ep}{1+e\sin\theta}$

opens down.

Or, if these are the equations of an ellipse, then the major axis is parallel to the x-axis

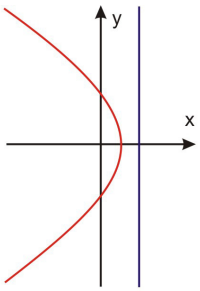


Denominator contains 1 - e sinθ



Denominator contains 1 + e sinθ

$r = \frac{ep}{1+e \cos \theta}$ opens to the left or vertically or major axis is the y-axis



Denominator contains 1 + e cosθ

If the polar equation is given with a constant in the numerator and with a sum of a constant and a term that contains either sin or cos in the denominator, the polar equation is a conic.

To graph such a polar equation these are steps to follow:

1. First determine which conic will match the given polar equation. To do this: Be sure the constant in the denominator is Equation A is equal to 1. If it is not, then divide the three terms in Equation A by this constant. Once the constant is ONE, the coefficient of the trig term is equal to e. If e > 1, the conic section is a hyperbola. If e = 1, the conic is a parabola and if e < 1, then the conic section is an ellipse.

When e = 0, the conic is a circle.

2. The numerator is the product of e and p (the distance between the focus and the directrix). Substituting the value of e, from step (1), p can now be found.

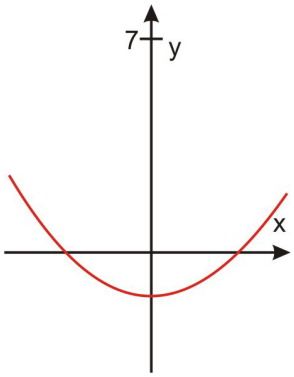
3. Examine the four general conic polar equations to see how the conic will open. This is determined by the sign of the trig term and whether the trig term is sin or cos.

Example: Graph the polar equation: $r = \frac{9}{3-3 \sin \theta}$

1. Divide all terms by 3 to put into the form $1 - e \sin \theta$ or $r = \frac{3}{1- \sin \theta}$

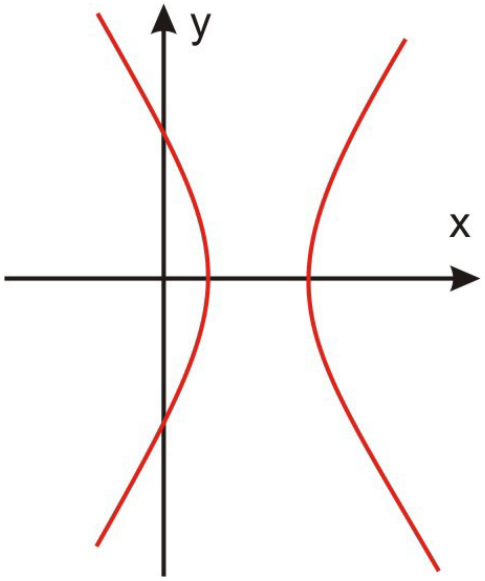
2. This has -e sin in denominator- therefore it opens up, e = 1 the conic is a parabola

3. ep = 3 and e = 1, p = 3



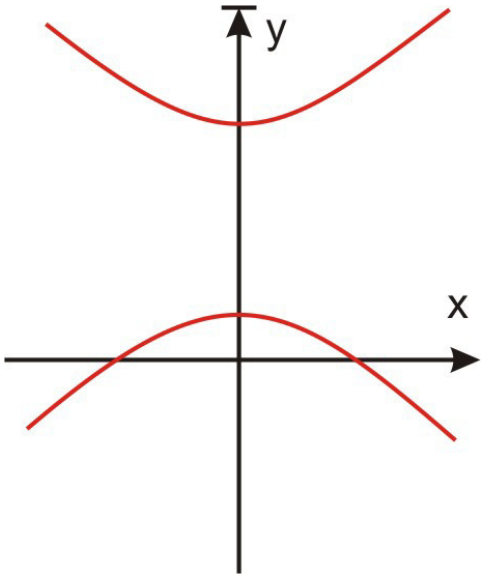
Example: Graph the polar equation: $r = \frac{3}{1+2\cos\theta}$

1. Since the constant in the numerator is already 1, no need to divide all terms.
2. e = 2, the coefficient of the cosine term. Therefore the conic is a hyperbola ep = 3, p = 1.5
3. This has a +2cos in denominator, so the conic opens left or horizontally.



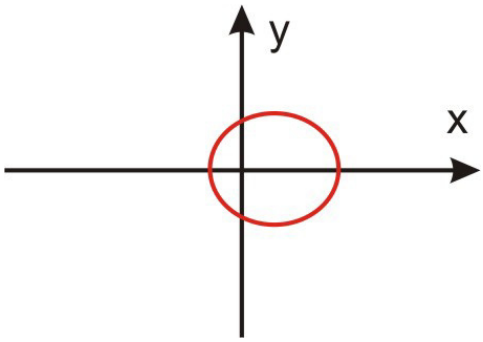
Example: Graph the polar equation: $r = \frac{5}{2+3\sin\theta}$

1. divide all terms by 2: $r = \frac{5/2}{1+(3/2)\sin\theta}$
2. e = 3/2, hyperbola., ep = 5/2, p = 5/3
3. 1 + 3/2sin in denominator- opens vertically



Example: Graph $r = \frac{2}{2-\cos\theta}$

1. divide all terms by 2: $r = \frac{1}{1-\frac{1}{2}\cos\theta}$
2. $e = \frac{1}{2}$, the conic an ellipse, ep = 1 or p = 2
3. the denominator contains $1 - \frac{1}{2}\cos\theta$, major axis is horizontal



Applications, Technological Tools

Technology note: Some graphing calculators can convert polar points to rectangular form and rectangular coordinates to polar form.

On the TI-83/84: go to [**ANGLE**](or [**2nd**]function) [**APPS**]. Scroll down to 5 or "R-Pr(" and press [**Enter**]. Next, enter the rectangular coordinates and close the parenthesis. Press [**Enter**], the "r" value appears. Scroll down to 6R-P , and the polar angle appears in decimal radian form.

Also under the [**ANGLE**] menu, commands 7 and 8 allow transformation from polar form to rectangular form.

Lesson Summary

There are at least two coordinate systems- rectangular and polar. Each can be used to graph a variety of equations. Additionally, once a point or an equation is graphed in one system, these points and equations can be transformed into the other system.

Points to Consider

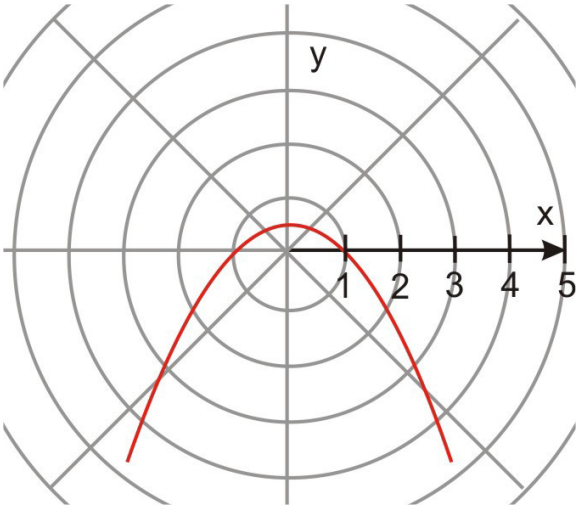
Sometimes an equation written in rectangular form, it can reveal characteristics that are hidden, unless transformed into polar. The same can be true for polar equations when transformed into rectangular form. The key here is to be able to transform one type into the other with ease, so that all characteristics of the equation can be known.

Review Questions

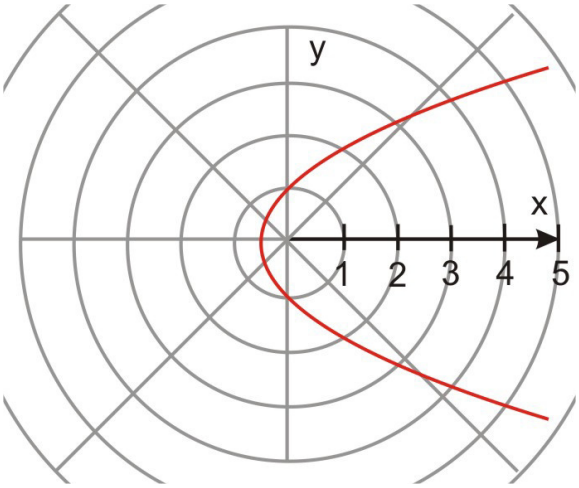
1. Convert the following: Convert the following rectangular coordinates to polar coordinates
   a. $(3, 3\sqrt{3})$
   b. $(-2, 2)$
   Convert the following polar coordinates to rectangular coordinates:
   c. $\left(4, \frac{2\pi}{3}\right)$
   d. $\left(-1, \frac{5\pi}{6}\right)$
2. Identify the conic section that each polar equation represents. Use the three steps to graph each polar equation: a. $r = \frac{1}{1+\sin\theta}$
   b. $r = \frac{1}{1-\cos\theta}$
   c. $r = \frac{6}{3-2\cos\theta}$
   d. $r = \frac{12}{5+3\cos\theta}$
   e. $r = \frac{8}{2-4\cos\theta}$

Review Answers

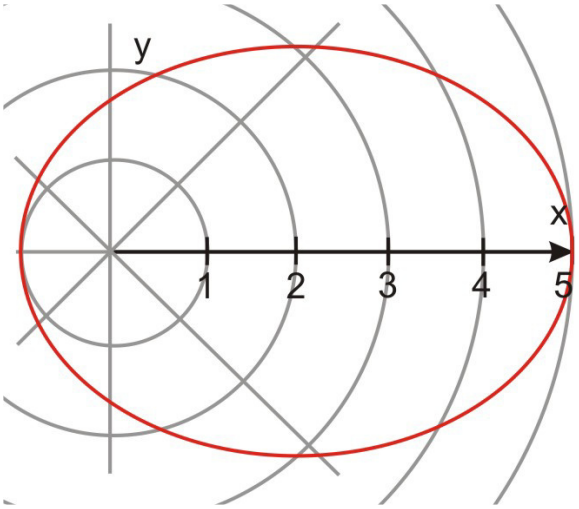
1. a. $(6, 66°)$b. $(2\sqrt{2}, 225°)$
   c. $(-2, 2\sqrt{3})$
   d. $\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$
2. a. parabola e = 1, + sin $\rightarrow$ opens down



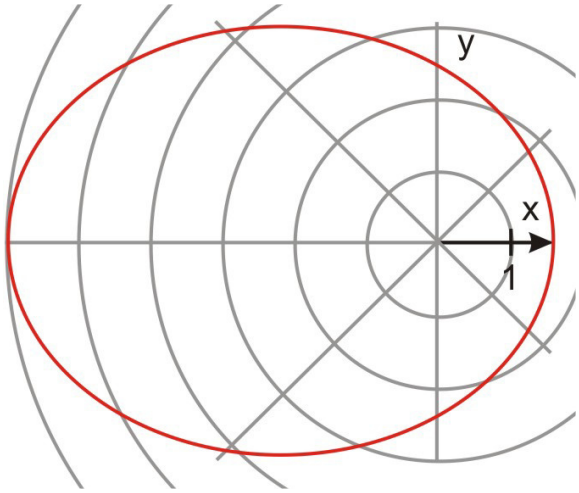
b. parabola e = 1, - cos $\rightarrow$ opens to the right



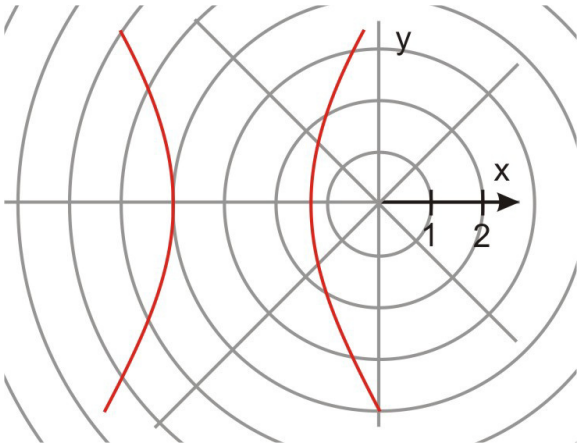
c. ellipse $e = \frac{2}{3} < 1$



d. ellipse $e = \frac{3}{5} < 1$



e. hyperbola, $e = 2 > 1$



5.3 Systems of Polar Equations

Learning Objectives

- Solve systems of polar equations graphically, numerically, and analytically
- Find all intersections of two polar equations graphed on the same set of axis
- Plot polar functions using a table of values
- Recognize equivalent polar equations

Introduction

Polar equations can be graphed using polar coordinates. Graphing two polar equations on the same set of axis may result in having point(s) of intersection.

All points on a polar graph are coordinates that make the equation valid. The coordinates of point(s) of intersection when substituted into each equation will make both of the equations valid.

Graph and Calculate Intersections of Polar Curves

One method to find point(s) of intersection for two polar graphs is by setting the equations equal to each other. Call the first equation r_1 and the second equation r_2 .Points of intersection are when $r_1 = r_2$, so set the equations equal and then solve the resulting trigonometric equation.

Example: Find the intersection of $r_1 = 3 \sin\theta \qquad r_2 = \sqrt{3} \cos\theta$

First set them equal

$3 \sin\theta = \sqrt{3} \cos\theta$

divide both sides of the equation by cos and 3:

$\frac{3 \sin\theta}{3 \cos\theta} = \frac{\sqrt{3} \cos\theta}{3 \cos\theta}$

and this simplifies to

$\frac{\sin\theta}{\cos\theta} = \frac{\sqrt{3}}{3}$

and then use the identity $\frac{\sin\theta}{\cos\theta} = \tan\theta$

$\tan\theta = \frac{\sqrt{3}}{3}$

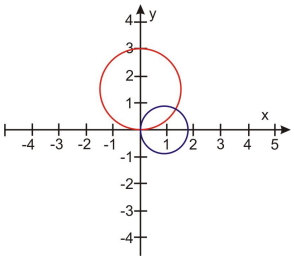
$\theta = \frac{\pi}{6}$ or $\frac{7\pi}{6}$

substitute $\frac{\pi}{6}$ in either equation to obtain r = 1.5

substitute $\frac{7\pi}{6}$ in either equation to obtain -1.5

***Note*:** the coordinates $\left(1.5, \frac{\pi}{6}\right)$ and $\left(-1.5, \frac{7\pi}{6}\right)$ represent the SAME polar point so there is only one solution to this equation.

Are we done? If we look at the graphs of r_1 and r_2, there is another point of intersection



when = 0,

$r_1 = 3 \sin = 3 \sin(0) = 0$

that is $r_1 = 3 \sin$ goes through the pole (0, 0).

For r_2 when $\theta = \frac{\pi}{2}$, $r_2 = 0$ that is $r_2 = \sqrt{3} \cos\theta$ goes through the point $\left(0, \frac{\pi}{2}\right)$. Therefore, both graphs go through the pole and the pole is a point of intersection.

The pole was NOT revealed as a point of intersection using the first step! (Why? Hint: How many ways are there to represent the pole in polar coordinates?) This shows us that after you use algebraic methods to find intersections at points other than the pole, you should also check for intersections at the pole.

Example: Find the point(s) of intersection for the two graphs:

$r_1 = 1$

$r_2 = 2 \sin 2$

1. First set $r_1 = r_2$ and solve:

$1 = 2 \sin 2\theta$

$\frac{1}{2} = \sin 2\theta$

Now, use a substitution $= 2$ to solve

$\frac{1}{2} = \sin \alpha$

$\alpha = \frac{\pi}{6}, \frac{5\pi}{6}$

Since $= 2$, solving for gives us

$\theta = \frac{\pi}{12}, \frac{5\pi}{12}$

But, recall that has the range 0 2 . Since we solved with 0 2 we actually need to consider values of with 0 4 . Why? Recall that sin(2) has two cycles between 0 and 2 , and so we add two more solutions,

$\alpha = \frac{13\pi}{6}, \frac{17\pi}{6}$

and since $= 2$,

$\theta = \frac{13\pi}{12}, \frac{17\pi}{12}$

Finally, we need to consider solutions when r = -1 because $r = 1$ and r = -1 are the same polar equation. So, solving

$-\frac{1}{2} = \sin \alpha$

$\alpha = \frac{7\pi}{6}, \frac{11\pi}{6}$

Again, using $= 2$ and adding solutions for the repetition gives us four more solutions,

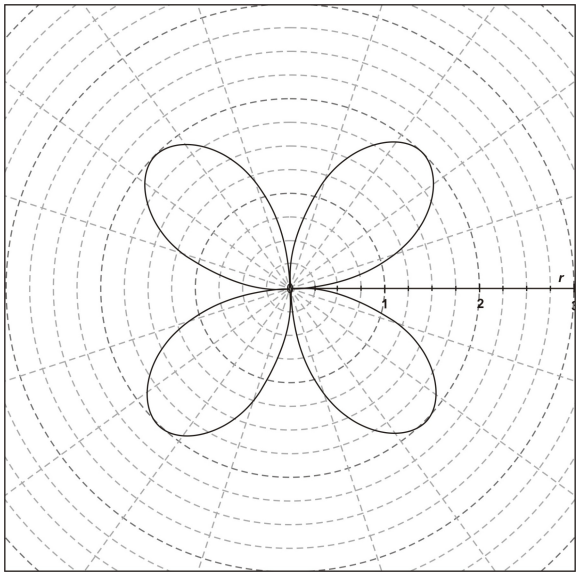
$\theta = \frac{7\pi}{12}, \frac{11\pi}{12}, \frac{19\pi}{12}, \frac{21\pi}{12}$

So in total, there are eight solutions to this set of equations.

Note: Recall-when solving trigonometric equations where the angle is requires looking at all potential values between 0 and 2 . When the angle is 2 as it is in this case, be sure to look for all potential values between 0 and 4 . When the angle is 3 as it is in this case, be sure to look for all potential values between 0 and 6 ., and so on.

2. Since r_1 cannot equal 0, the pole is not on its graph and not a point of intersection.

3. The graph reveals eight points of intersection which were found in step (1).



Equivalent Polar Curves

Recall from trigonometry that the same trigonometric function can often be written in multiple ways. For example, above you worked with $r = 2 \sin(2\ )$. Recall the identity

$\sin(2\ ) = 2 \sin \quad \cos \ .$

Substituting that identity into the equation yields

$r = 2 \sin(2\ )$

$r = 2 (2 \sin \quad \cos \ )$

$r = 4 \sin \quad \cos$

Thus, the two equations

$r = 2 \sin(2\ )$ and $r = 4 \sin \quad \cos$ are equivalent. To verify this, try graphing both equations on the same axes.

We also discovered another interesting issue with polar graphing above: the equation $r = 1$ makes a circle of radius 1 centered on the origin, and so does the equation $r = -1$. Thus $r = 1$ is the same curve as $r = -1$. In general $r = k$ is equivalent to $r = -k$.

Both of these examples highlight that polar equations can be tricky. When you solve problems involving polar equations, you need to consider ALL solutions to the equation, not just the *immediate* solutions from your first attempt at solving. Graphing polar equations is a good way to check for all solutions.

Given Two Polar Curves, Find All Intersection Points

Example: Find point(s) of intersection, if any exist, for the following pair of equations:

a. $r_1 = 2$ and $r_2 = \sec$

Here we will use a table of values for each function, solving by quadrant. Recall that the period of sec is $2\ $.

For the first quadrant

Table 5.1:

(angle)	0	/6	/4	/3	/2
r_1 (distance)	2	2	2	2	2
r_2	1	1.15	1.4	2	und

For the second quadrant

Table 5.2:

(angle)	2 /3	3 /4	5 /6	
r_1 (distance)	2	2	2	2
r_2	-2	-1.4	-1.15	-1

For the third quadrant:

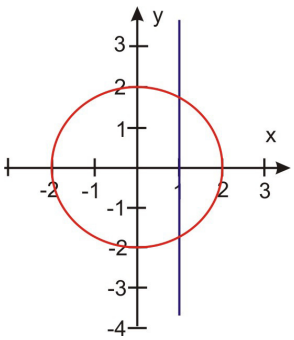
Table 5.3:

(angle)	7 /6	5 /4	4 /3	3 /2
r_1 (distance)	2	2	2	2
r_2	-1.15	-1.4	-2	und

For the fourth quadrant

Table 5.4:

(angle)	5 /3	7 /4	11 /6	2
r_1 (distance)	2	2	2	2
r_2	2	1.4	1.15	1



Note that the 3rd and 4th quadrant repeat 1st and 2nd quadrant values.

Observe in the table of values that (2, /3) and (2, 5 /3) are the points of intersection. Look at the curves- the first equation yields a circle while the second yields a line. The maximum number of intersecting points for a line and a circle is 2. The two points have been found.

Another way to solve polar equations analytically is to set both of the equations equal to each other, that is let $r_1 = r_2$, or equivalently

2 = sec .

Using rules from trigonometry this equation can be written as

½ = cos ,

by taking the reciprocal of both sides.

Solving for cos = ½ , = /3 or 5 /3. These are the same values for the table of values yield. Point(s) of intersection for Polar curves drawn on the same axis can be found by equating their equations and then solving using the methods for solving trigonometric equations.

Example: Find the point(s) of intersection for these pair of Polar equations:

b. $r = 2 + 4 \sin$ and $= 60°$

The equation = 60° is a line making a 60° angle with the r axis. Now we make a table of values for $r = 2 + 4 \sin$

For the first quadrant:

Table 5.5:

(angle)	0	30	45	60	90
R (distance)	2	4	4.83	5.46	6

For the second quadrant:

Table 5.6:

(angle)	120	135	150	180
R (distance)	5.46	4.83	4	2

For the third quadrant:

Table 5.7:

(angle)	210	225	240	270
R (distance)	0	-.83	-1.46	-2

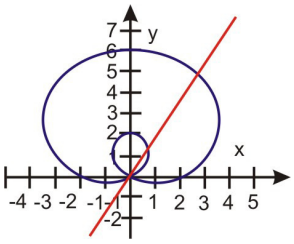
For the fourth quadrant

Table 5.8:

(angle)	300	315	330	360
R (distance)	-1.46	-.83	0	2

Notice there are two solutions in the table., (60, 5.46) and (240, -1.46) = (60, 1.46). Recall that when $r <$ 0, you plot a point (r,), by rotating 180° (*or*).

Finally, we need to check the pole: $r = 2 + 4 \sin$ passes through the pole for $= 330°$, and $= 60°$ also passes through the pole. Thus, the third point of intersection is (0, 0).



Example: $r_1 = 2 \cos$ and $r_2 = 1$.

Here we make a table:

For the first quadrant:

Table 5.9:

(angle)	0	/6	/4	/3	/2
r_1 (distance)	2	$\sqrt{3}$	$\sqrt{2}$	1	0
r_2	1	1	1	1	1

For the second quadrant:

Table 5.10:

(angle)	2 /3	3 /4	5 /6	
r_1 (distance)	-1	$-\sqrt{2}$	$-\sqrt{3}$	-2
r_2	1	1	1	1

For the third quadrant:

Table 5.11:

(angle)	7 /6	5 /4	4 /3	3 /2
r_1 (distance)	$-\sqrt{3}$	$-\sqrt{2}$	-1	0
r_2	1	1	1	1

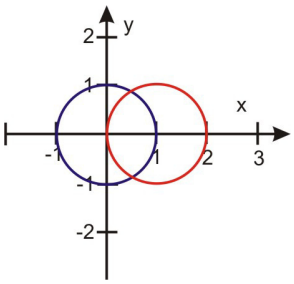
For the fourth quadrant:

Table 5.12:

(angle)	5 /3	7 /4	11 /6	2
r_1 (distance)	1	$\sqrt{2}$	$\sqrt{3}$	2
r_2	1	1	1	1

So the unique solutions are at $\theta = \frac{\pi}{3}, \frac{4\pi}{3}$. There are also two repeated solutions in this set (can you find them?).

Here is a graph showing the two solutions:



Applications, Technological Tools

To graph a Polar equation using a TI calculator:

Press: "**[MODE]**" On the 4th line be sure to Press "**[POL]**" (for Polar)

Then Press "**[Y=]**" Note: the "y" has now been replaced with "r". Input the equation. ***Note***: When in polar mode, pressing "X, T, , n" key, is the default

Note: There may be a need to adjust, as always, the window values, in order to see the graph within the constraints of the calculator window.

Lesson Summary

1. One method for finding point(s) if intersection for two graphs is to plot points for both graphs on the same axis. A second method, is to let the value for r_1 equation equal value for r_2 .

The procedure for finding the point or points of intersection for two polar graphs is almost the same as it is for rectangular graphs.

2. The pole can be represented by many different polar coordinates (example: (0, 0) or (0, 90º) or $\left(0 \ \frac{-\pi}{4}, \right)$, it is important to check if both graphs go through the pole.

3. Analytically graph the two equations on the same polar axis to check if all of the point(s) of intersection have been found in steps (1) and (2).

Points to Consider

The rectangular equation $x = 2$ and the polar equation $r = 2 \csc$, reveal the same curve. Yet in this example, the rectangular form has only one variable in the equation, whereas the polar form has two variables. Although the polar form may be useful at times, the rectangular form is easier to work with.

$x^2 + y^2 = 1$ and r = 1 also reveal the same curve. Here, the polar form is more simplified since it only contains one variable. What this reveals is that sometimes one form of an equation may be more useful in terms of simplicity to work with, than the other form.

Being able to recognize that two apparently different equations may reveal the same curve can be difficult at first, but having this knowledge provides two potentially useful ways to work with a problem. Either form can be used, depending on which provides the simplest or fastest method. Furthermore, having two equivalent methods offers a unified understanding of the mathematics involved which can provide insights that may not be achieved when thinking within the context of any one individual system.

Review Questions

1. Find the Point(s) of intersection for each system of equations. Graph to verify your solution.

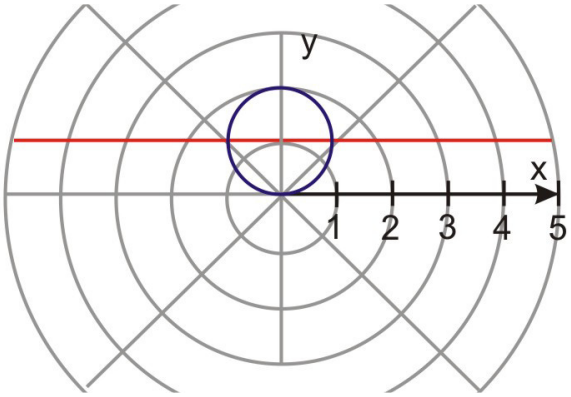
Table 5.13:

a. $r_1 = \csc$	$r_2 = 2 \sin$
b. $r_1 = \cos$	$r_2 = 1 + \sin$
c. $r_1 = \sin$	$r_2 = \sin 2$
d. $r_1 = -4 \sin$	$r_2 = -4 \cos$
e. $r_1 = 1 - 2 \sin$	$r_2 = \sqrt{9 \cos(\theta)}$
f. $r_1 = 1 - \cos$	$r = 4 \cos(3)$

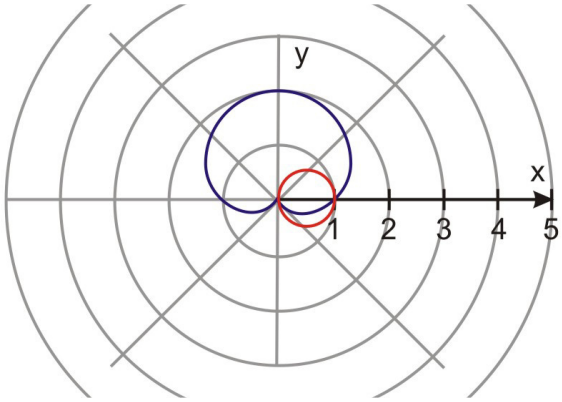
2.

Review Answers

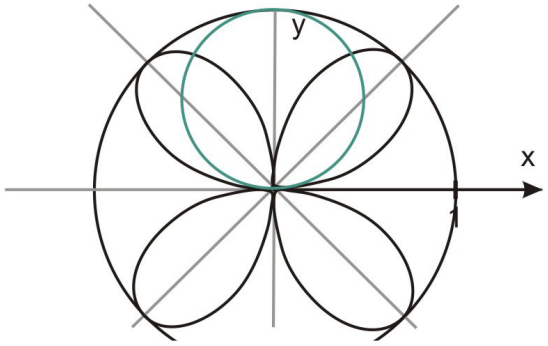
1. a. $\left(\sqrt{2}, \frac{\pi}{4}\right)$, $\left(\sqrt{2}, \frac{3\pi}{4}\right)$



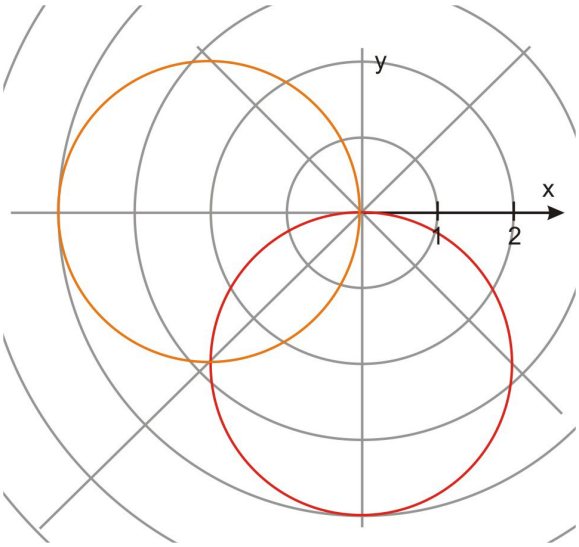
b. (0, 0), (1, 0)



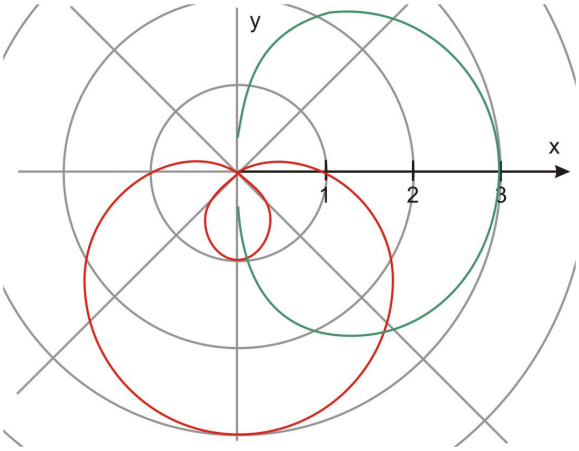
c. (0, 0), $\left(\frac{\sqrt{3}}{2}, \frac{2\pi}{3}\right)$, $\left(\frac{\sqrt{3}}{2}, \frac{\pi}{3}\right)$



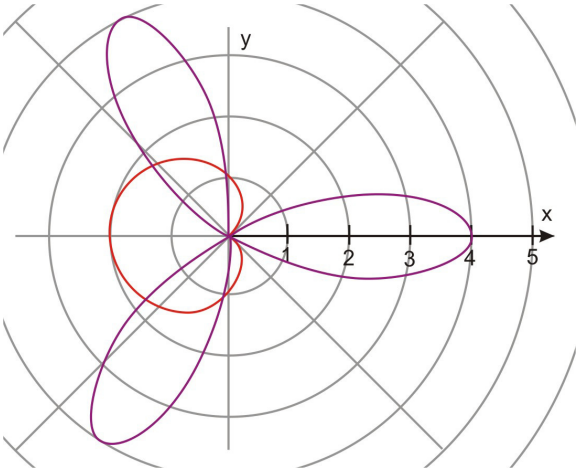
d. (0, 0), $\left(2\sqrt{2}, \frac{5\pi}{4}\right)$



e. (1, 276°), (2.44, 313°)



f. (0, 0), (1.08, 95°), (1.77, 142°), (1.77, 218°), (1.08, 265°)



5.4 Imaginary and Complex Numbers

Learning Objectives

- Write square roots with negative radicands in terms of i
- Recognize and write complex numbers in standard form
- Describe the relationship between the sets of integers, rational numbers, real numbers and complex numbers
- Plot $z = a + bi$ in the complex number plane

Introduction

While working with quadratic equations, you may have solved an equation such as:

$(x - 1)^2 + 4 = 0$.

No matter which method of solving quadratics you used, the solutions to that equation are not real numbers, and you find that they are $1 + 2i$ and $1 - 2i$. (Recall that $\sqrt{-1} = i$). These solutions combine **imaginary** and **real** numbers, and are called **complex** numbers.

The use of the word *imaginary* does not mean these numbers are useless. For a long period in the history of mathematics, it was thought that the square root of a negative number was in fact only within the mathematical imagination without real-world significance hence, imaginary. That has changed. Mathematicians now consider the imaginary number as another set of numbers that have real significance, but do not fit on what is called the number line—and engineers, scientists, and others solve real world problems using complex numbers!

Recognize

Recognize $i = \sqrt{-1}$, $\sqrt{-x} = i\sqrt{x}$

Where did complex numbers come from? If you solve the equation $x^2 + 1 = 0$, you get $x = \pm\sqrt{-1}$. But there is no real number that, when multiplied by itself, yields -1. To fix this problem, mathematicians defined the imaginary constant, i. by definition,

$i = \sqrt{-1}$

or squaring both sides,

$i^2 = -1$

Recall that you can simplify radicals by factoring out perfect squares in the radicand. For instance, $\sqrt{8} = \sqrt{4 \cdot 2} = \sqrt{4}\sqrt{2} = 2\sqrt{2}$. The same procedure works with i. If you have a negative number in the radicand, you can factor out the -1 and use the identity $i = \sqrt{-1}$ to simplify.

Example: Simplify $\sqrt{-5}$

Solution:

$\sqrt{-5} = \sqrt{(-1) \cdot (5)}$

$= \sqrt{-1}\sqrt{5}$

$= i\sqrt{5}$

This also works in combination with the other method of factoring out perfect squares. See the following example.

Example: Simplify $\sqrt{-72}$

Solution:

$\sqrt{-72} = \sqrt{(-1) \cdot (72)}$

$= \sqrt{-1}\sqrt{72}$

$= i\sqrt{72}$

But, we're not done yet. $72 = 36 \cdot 2$, so

$i\sqrt{72} = i\sqrt{36}\sqrt{2}$

$= i(6)\sqrt{2}$

$= 6i\sqrt{2}$

Standard Form of Complex Numbers (a + bi)

Sometimes when you solve a quadratic equation, the solution has both a real part and an imaginary part. For example, if you want to solve

$(x-1)^2 + 4 = 0$

then

$(x-1)^2 = -4$

$x - 1 = \pm\sqrt{-4}$

$x - 1 = \pm\sqrt{-1}\sqrt{4}$

$x - 1 = \pm 2i$

$x = 1 \pm 2i$

$x = 1 \pm 2i$ *or* $1 - 2i$

Notice that these two solutions involve a real part, 1, and an imaginary part, $\pm 2i$

$z = a + bi$ is the **standard** or **rectangular** form of a complex number.

A complex number is a number that has a real part (in this case a) and an imaginary part, that is, the imaginary number i with a coefficient b.

Set of Complex Numbers (complex, real, irrational, rational, etc.)

The complex numbers are a superset of the real numbers. Given $z = a + bi$, if $b = 0$ then z is a real number. Every real number can be written as a complex number (just let it equal a), but there are many more complex numbers than real numbers. Hence the complex numbers are a superset of the real numbers.

When you were first introduced to mathematics, you probably used **positive whole numbers**, that is 1, 2, 3, 4,...

Later, negative whole numbers are investigated. The set of all whole numbers, both positive and negative, including the number zero, is known as **integers**: ... - 2, - 1, 0, 1, 2,...

Later, students are introduced to fractions. The set of all numbers that CAN be expressed as a quotient of two integers (where the denominator is not zero) is called the set of **Rational Numbers**. Rational Numbers can also be expressed as a terminating or repeating decimal. Some rational numbers are $-1, \frac{3}{5}, -\frac{7}{3}, 1,000,002, 0$. Of course there are an infinite number of rational numbers between any two whole numbers, so listing all rational numbers neatly is difficult (but it is possible–can you think of a way to do

it?).

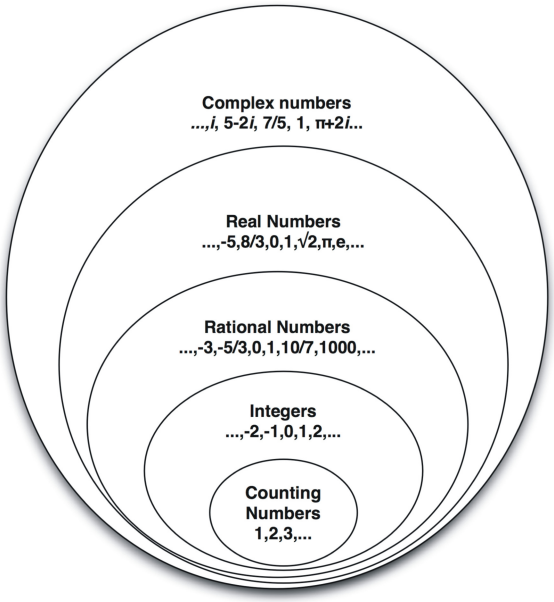
Notice, that all integers are in the set of rational numbers (for example, 5 CAN be written as the quotient of 10 and 2 since $5 = \frac{10}{2}$), so the integers are a subset of the rational numbers. Finally, when working with circles students encounter a number that can be approximated as a quotient of two integers but cannot be expressed EXACTLY as that quotient, that is the number . Recall that was often expressed as APPROXIMATELY $\frac{22}{7}$ or 3.14, BUT NOT EXACTLY THOSE VALUES.

When first exploring using the Pythagorean Theorem to find the length of a diagonal of a square whose side is 1, the number $\sqrt{2}$ was introduced. $\sqrt{2}$ often was approximated as 1.4 or 1.414, but again you can't possibly write out all of the decimals in $\sqrt{2}$. These two numbers are examples of IRRATIONAL numbers, that is numbers that cannot be expressed as a quotient of two integers, and therefore CANNOT be expressed as a terminating or repeating decimal. The set of all rational and irrational numbers together is called REAL numbers.

Finally, when working again with the Pythagorean Theorem in the coordinate plane (where "negative distances" are possible), negative values appeared within the square root! But what number times itself can result in a negative number?

Historically, when this occurred, mathematicians thought that this was only an oddity of the theorem and not something that can actually exist. They therefore called such numbers imaginary. But, some real-world problems can be solved with imaginary numbers.

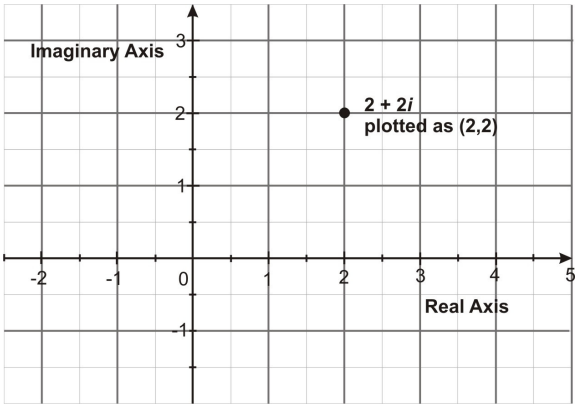
The set of all real numbers AND imaginary numbers is called the set of Complex numbers.



Complex Number Plane

In standard form $z = a + bi$, a complex number can be graphed using rectangular coordinates (a, b). a represents the x - coordinate, while b represents the y - coordinate. Alternatively, the x - coordinate represents "real number" values, while the y - coordinate represents the "imaginary" values.

For example, given the complex number in standard form: $z = 2 + 2i$, you can graph this number in the coordinate plane To graph this point, the coordinate (2, 2) is graphed as shown below:



Lesson Summary

When graphing a complex number using rectangular coordinates, the x-axis plots the real number, while the y-axis plots the coefficient of the imaginary number.

Points to Consider

Given a point in a rectangular coordinate system that represents a complex number, multiply that complex number by i and graph this new complex number. If the points that represent the original complex number and the new complex number have a line drawn from the origin to each point- note the angle between the two lines. Multiply the second complex number by i and plot this third point. Draw a line from the origin to this point. Note the angle between the second line and the third line. What appears to happen when a complex number is multiplied by i?

Review Questions

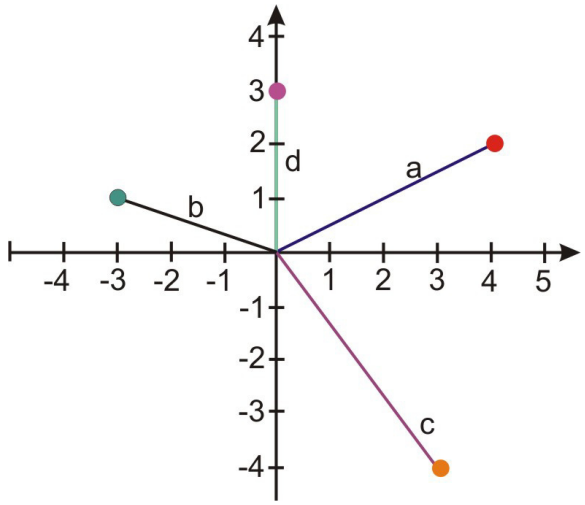
Using rectangular coordinate system, graph

1. Simplify the following radicals a. $\sqrt{-9}$
   b. $\sqrt{-12}$
   c. $\sqrt{-17}$
   d. $\sqrt{140-108}$
2. Solve each equation and express it as a complex number. (Note: If the imaginary part is 0, you can still express the solution as $a + bi$, but you will have $b = 0$a. $x^2 + 24 = 0$
   b. $2x^2$ - $4x + 7 = 0$
3. Plot each of the following complex numbers: a. $(4 + 2i)$
   b. $(-3 + i)$
   c. $(3 - 4i)$
   d. $3i$

Review Answers

1. a. $3i$b. $2i\sqrt{3}$
   c. $i\sqrt{17}$
   d. $\sqrt{32} = 6\sqrt{2}$

2. a. $x = \pm 2\sqrt{6}i$ b. $x = 1 \pm 2\sqrt{3}i$
3.



5.5 Operations on Complex Numbers

Learning Objectives

- Work with roots of quadratic equations by using the discriminant
- Add and subtract complex numbers
- Find the complex conjugate of a complex number
- Multiply and divide complex numbers
- Use your calculator to add, subtract, multiply, and divide complex numbers

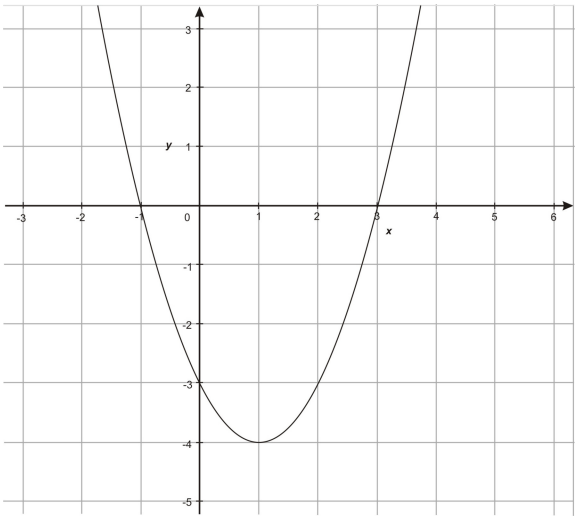
Quadratic Formula

If $ax^2 + bx + c = 0$

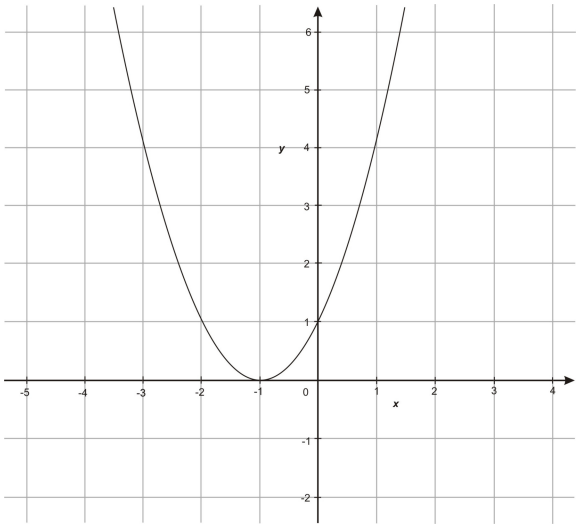
then $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Recall that b^2 - $4ac$ is called the discriminant.

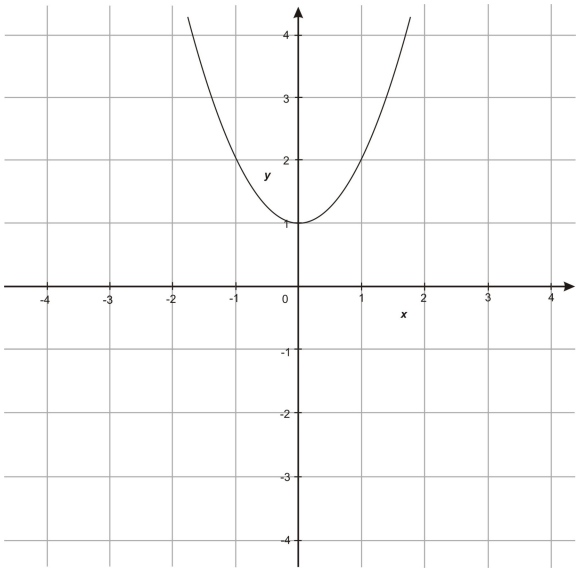
If b^2 - $4ac > 0$ then there are two unequal real solutions.



If b^2 - $4ac = 0$ then there are two equal real solutions.



If b^2 - $4ac < 0$ then there are two unequal complex solutions.



Example: Solve $x^2 + 4x + 6 = 0$.

$x^2 + 4x + 6 = 0$

$a = 1, b = 4, c = 6$

$x = \frac{-4 \pm \sqrt{(4)^2 - (4)(1)(6)}}{2(1)}$

Since b^2 - $4ac$ = -8, the solutions will be complex

$x = \frac{-4 \pm \sqrt{-8}}{2}$

or $x = \frac{-4 \pm 2i\sqrt{2}}{2}$

$x = -2 \pm i\sqrt{2}$

Sums and Differences of Complex Numbers

Recall that a complex number is given in the form $a + bi$ where a and b are real numbers and i is the imaginary constant, $i^2 = -1$.

When adding (or subtracting) two or more complex numbers the fastest method will be to add (or subtract) the real components to obtain the sum of the real numbers, and then separately add (or subtract) the imaginary coefficients to obtain the sum of the imaginary numbers or:

$(a + bi) + (c + di) = [a + c] + [b + d]i$

Example: Perform the Indicated Operation:

$(6 + 3i) + (-5 + 2i) = (6 + -5) + (3 + 2)\ i = 1 + 5i$

$(3 - 2i) - (2 - 4i) = (3 - 2) + (-2 - (-4))\ i = 1 + 2i$

$(6) + (4 - 3i) = (6 + 4) + (0 + (-3))\ i = 10 - 3i$

Products and Quotients of Complex Numbers (conjugates)

Multiplying Complex Numbers:

When multiplying complex numbers in rectangular form, recall the method for multiplying two binomials (sometimes called FOIL): $(m + n)(x + y) = mx + my + nx + ny$. We use the same procedure for multiplying complex numbers:

$(a + bi) + (c + di) = ac + adi + bci + bdi^2$

But, unlike the algebraic expression, the above expression contains the number, i.

Recall that $i^2 = -1$, so $bdi^2 = bd(-1) = -bd$ and

$adi + bci$ can be combined and then factored as $(ad + bc)i$. Thus we have the general result,

$(a + bi) + (c + di) = (ac - bd) + (ad + bd)i$

Example: Multiply: $(6 + 3i)(2 - 3i)$

$(6 + 3i)(2 - 3i) = 12 - 18i + 6i - 9i^2$

$= 21 - 12i$

(Combining like terms $-9i^2$ reduces to -9(-1) or 9)

Example: Multiply: $(5 - 7i)(5 + 7i)$

$(5 - 7i)(5 + 7i) = 25 + 35i - 35i - 49i^2$

or

$= 25 + 0 - 49(-1)$

$= 74$

Note: The product of the complex number (5 - 7i) and the complex number (5 + 7i) is a real number. When a complex number is multiplied by another complex number to produce a real number, the two complex numbers are called **complex conjugates.**

Example: Multiply $i(6 - 2i)$

$6i - 2i^2 = 6i - 2(-1)$

$= 6i + 2$

But, the real part of a complex number is generally written first, so we can write this as

$= 2 + 6i$

Dividing Complex Numbers

To divide two complex numbers is similar to dividing two irrational numbers. Recall that in that problem, the procedure was to find the irrational **conjugate** of the denominator and then multiply both the numerator and the denominator by this conjugate

Divide: $\frac{3}{1+\sqrt{2}}$

First find the irrational conjugate of the denominator: $1 - \sqrt{2}$, then multiply both the numerator and the denominator by this value:

$\frac{3}{1+\sqrt{2}} \times \frac{1-\sqrt{2}}{1-\sqrt{2}} = \frac{3-3\sqrt{2}}{1-2}$

this reduces to

$= \frac{3-3\sqrt{2}}{-1}$

or

$= -3 + 3\sqrt{2}$

In this case since you are interested in eliminating the complex numbers from the denominator, find the **complex conjugate** of the denominator and multiply BOTH the numerator AND the denominator by it.

You find the complex conjugate in the same way you found the conjugate of irrational numbers, change the sign of the imaginary part. For instance, the complex conjugate of 4 + 3i is 4 – 3i

A complex number multiplied by its complex conjugate will yield a real number. By recalling $(a + b)(a - b) = a^2 - b^2$ the complex conjugate can be found:

The conjugate of 4 + 3i is found by retaining the real part (4) and reversing only the sign of the imaginary part (that is, 3i becomes -3i)
$(4 + 3i)(4 - 3i) = 16 - 12i + 12i - 9i^2$ Notice that $-12i$ and $12i$ cancel. Also recall that $i^2 = -1$
$16 + 9 = 25$ $(4 + 3i)(4 – 3i) = 25$

The product of this complex number and its conjugate is 25.

When multiplying complex numbers sometimes intuition about the nature of the product can mislead.

For example in (a + b)(a + b), where all of the terms are real numbers, no terms of each of the four products will cancel. Some of the terms may be combined. However in (1 + i)(1 - i), where some terms are real numbers and some terms are imaginary numbers, this is no longer true. Two of these terms cancel: the first product yields 1 while the last product yields i^2 or -1, and those terms cancel!

Example: Find the quotient: $\frac{6-3i}{4+3i}$

First, observe that the complex conjugate of the denominator is 4 – 3i

Multiply both the numerator and the denominator by 4 – 3i:

$\frac{6-3i}{4+3i} \times \frac{4-3i}{4-3i} = \frac{24-18i+12i+9i^2}{16-12i+12i-9i^2}$

$= \frac{15-6i}{25}$

$\frac{6-3i}{4+3i} = \frac{15-6i}{25}$

This number can also be written as $\frac{15}{25} - \frac{6i}{25}$ or $0.6 – 0.24i$

Applications, Technological Tools

A graphing calculator can perform operations with complex numbers. Press mode. Scroll down until: real $a + bi$ *re* "0302 i is seen, then select $a + bi$. Press Quit. Now the calculator is able to perform operations with complex numbers in $a + bi$ form.

When the calculator is in complex number mode, be sure to use parenthesis to group the parts of the complex numbers. For example, enter $1 + 3i$ as $(1 + 3i)$. Try doing some of the calculations from this section on your calculator to verify that complex mode works.

Lesson Summary

When adding and subtracting complex numbers add/subtract the real numbers and then add/subtract the imaginary numbers.

When multiplying complex numbers use the FOIL method of multiplication. Be sure to substitute $i^2 = -1$ when appropriate.

When dividing complex numbers, write the problem as a fraction and then multiply both numerator and denominator by the conjugate of the denominator. If this process is done successfully, there will be only a real number in the denominator.

Points to Consider

Complex numbers- that is numbers that have a real and an imaginary part, also known as components, have characteristics similar to working with two unlike terms. However, they also have major differences such as not having as a final answer to a question, a denominator with an imaginary component or being sure to reduce i^2 to -1.

Review Questions

Perform the indicated operations:

1. a. $(-8 - 2i) + (5 - 2i)$b. $(6 + i) - (5 - 2i)$
   c. $(-3i) - (2 - 3i)$
   d. $(3 + 2i)(5 - i)$
   e. $(0 + 4i)(3 - 1)$
   f. $(5 - 3i)^2$
   g. $\frac{6-i}{4+i}$
   h. $(5 - 2i) \div (-2 + 3i)$

Review Answers

1. a. $(3 - 4i)$b. $(1 + 3i)$
   c. $(-2 + 0i)$
   d. $(17 + 7i)$
   e. $(4 + 12i)$
   f. $(16 - 30i)$
   g. $\frac{23-10i}{17}$
   h. $\frac{-16-11i}{13}$

5.6 Trigonometric Form of Complex Numbers

Learning Objectives

- Convert complex numbers into rectangular and polar coordinates
- Use the polar form of complex numbers to explore powers of complex numbers

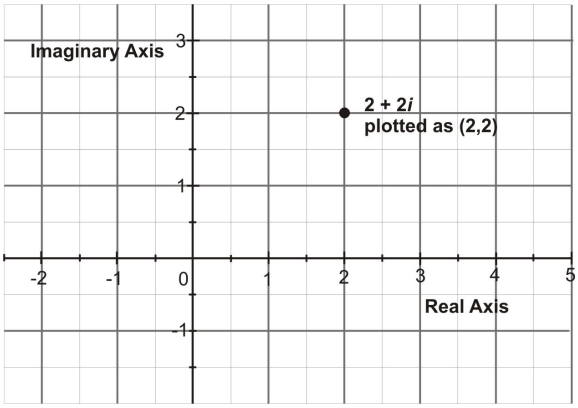
Introduction

You have learned that rectangular graphs can be put into polar form, and that points in rectangular coordinates can be plotted in the polar coordinate system. In this section you will learn how to do the same process with complex numbers.

Complex Plane

In standard form $z = a + bi$, a complex number can be graphed using rectangular coordinates (a, b). 'a' represents the x - coordinate, while 'b' represents the y - coordinate. Alternatively, the x - coordinate represents "real number" values, while the y - coordinate represents the "imaginary" values.

For example, given the complex number in standard form: $z = 2 + 2i$, you can graph this number in the coordinate plane To graph this point, the coordinate (2, 2) is graphed as shown below:



Relationships among x, y, r, and

You can also graph complex numbers in the polar coordinate system using the familiar substitutions.

If you are given r and , then use

$x = r \cos$ and $y = r \sin$

If you are given x and y, then

$r = \sqrt{x^2 + y^2}$ and $\tan \theta_{ref} = \left|\frac{y}{x}\right|$

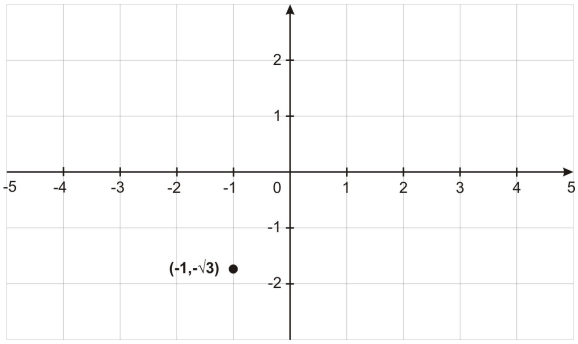
All the same rules and procedures for converting points represented by a real pairs of numbers in the rectangular plane apply to converting complex numbers into polar form. We review the steps for conversion below.

Trigonometric or Polar Form of a Complex Number (r cis)

In the example above, we graphed the complex number $z = a + bi$ in rectangular coordinate system. Recall that there is another coordinate system we can use, the polar coordinate system. In a previous lesson you learned that rectangular coordinates (x, y) can be transformed into polar coordinates $(r, \ )$.

Here we will use that basic conversion to rewrite $z = a + bi$ in another (sometimes more convenient) form that is based on the polar conversion. This new form is called the **trigonometric form** of a complex number.

Start with the complex number $z = -1 - i\sqrt{3}$. You can graph it in the rectangular coordinate system:



In Polar form, we find r with

$r = \sqrt{a^2 + b^2}$

$= \sqrt{(-1)^2 + (-\sqrt{3})^2}$

$= \sqrt{1 + 3}$

$= \sqrt{4}$

$= 2$

and to find ,

$\tan \theta_{ref} = \left|\frac{-\sqrt{3}}{-1}\right|$

$\tan \theta_{ref} = \sqrt{3}$

$\theta_{ref} = \tan^{-1} \sqrt{3}$

$\theta_{ref} = \frac{\pi}{3}$

Since this angle is in the 4^{th} quadrant, $\theta = \frac{4\pi}{3}$.

Now, there is a problem with using a conversion from rectangular form to polar form like

$a + bi \rightarrow (r, \theta)$

$-1 - i\sqrt{3} \rightarrow \left(2, \frac{4\pi}{3}\right)$

The problem is that we have lost the i. So, in order to "keep track" of the imaginary part, we can use another form. Recall that we used the substitutions

$x = r \cos$ and $y = r \sin$.

Using this fact, and the values we know for r and , we can write

$z = -1 - i\sqrt{3} = 2 \cos \frac{4\pi}{3} + 2\, i \sin \frac{4\pi}{3}$

Finally, factoring the 2,

$z = 2\left(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}\right)$

Our new form, $z = 2\left(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}\right)$ is in the form z = **r**(**c**os + **i**sin), where r and are the polar coordinates of the imaginary number. This form will be used very often in the future. It has an abbreviated form known as "rcis "

So, the complex number: $z = -1 - \sqrt{3}i$, the rectangular point $(-1, -\sqrt{3})$, the polar point: $\left(2, \frac{4\pi}{3}\right)$, and $2\left(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}\right)$ or 2 cis $\left(\frac{4\pi}{3}\right)$ all represent the same number.

Each of these forms has benefits in specific situations as we will show shortly.

Steps for Conversion

To convert from polar to rectangular form, the distance that the point (2, 2) is from the origin can be found by

$d = \sqrt{x^2 + y^2}$ or $\sqrt{2^2 + 2^2}$ $d = \sqrt{8}$ or $2\sqrt{2}$

The reference angle (i.e. the corresponding angle in the first quadrant) that the line segment between the point and the origin can be found by

$\tan \theta_{ref} = \left|\frac{y}{x}\right|$

for $z = 2 + 2i$,

$\tan \theta_{ref} = \frac{2}{2}$

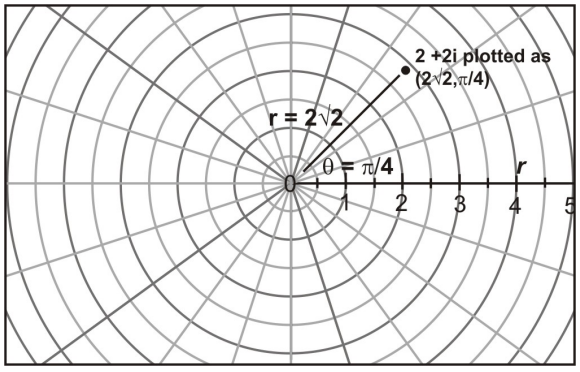
$\tan \theta_{ref} = 1$.

Since this point is in the first quadrant (both the x and y coordinate are positive) the angle must be 45° or $\frac{\pi}{4}$ radians.

It is also possible that when tan = 1 the angle can be in the third quadrant or $\frac{5\pi}{4}$ radians. But this angle will not satisfy the conditions of the problem, since a third quadrant angle must have both x and y as negatives.

Note: When using $\tan \theta = \frac{y}{x}$, you should first consider, the quotient $\left|\frac{y}{x}\right|$ and find the first quadrant angle that satisfies this condition. This angle will be called the **reference angle,** denoted θ_{ref}. Find the actual angle by analyzing which quadrant the angle must be given the x and y signs.

The complex number $2 + 2i$ or (2, 2) in rectangular form has polar coordinates $\left(2\sqrt{2}, \frac{\pi}{4}\right)$

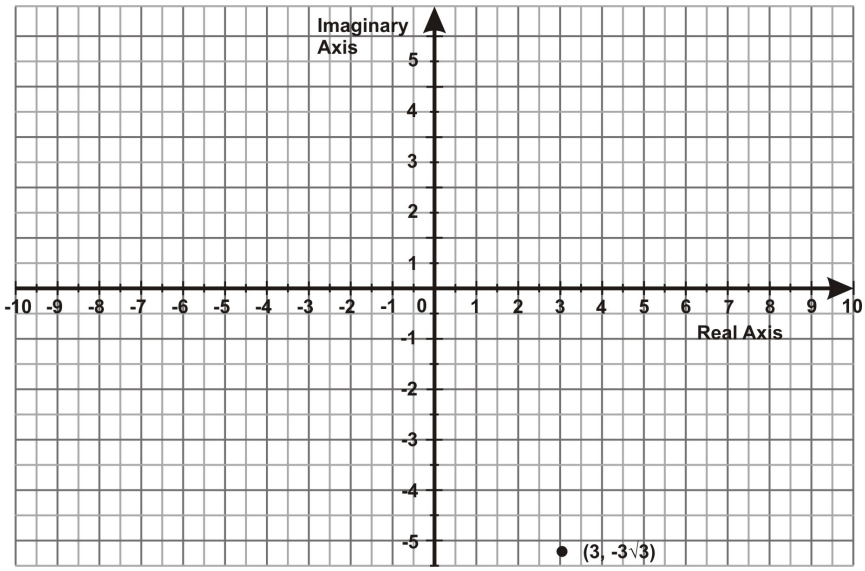


Example: Find the Polar coordinates that represent the complex number $z = 3 - 3\sqrt{3}i$

a = 3 and b = $-3\sqrt{3}$: the rectangular coordinates of the point is $\left(3, -3\sqrt{3}\right)$.

Now, draw a right triangle in standard form. Find the distance the point is from the origin and the angle

the line segment that represents this distance makes with the +x axis:



We know a = 3, $b = -3\sqrt{3}$

$r = \sqrt{3^2 + (-3\sqrt{3})^2}$

$= \sqrt{9 + 27}$

$= \sqrt{36}$

$= 6$

And for the angle,

$\tan \theta_{ref} = \left|\frac{(-3\sqrt{3})}{3}\right|$

$\tan \theta_{ref} = \sqrt{3}$

$\theta_{ref} = \frac{\pi}{3}$

But, since it is a 4^{th} quadrant angle

$\theta = \frac{5\pi}{3}$

The rectangular point $(3, -3\sqrt{3}i)$ is equivalent to the polar point $\left(6, \frac{5\pi}{3}\right)$.

In rcis form, $(3, -3\sqrt{3}i)$ is $6\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right)$.

Applications, Technological Tools

Some graphing calculators can convert polar to rectangular and rectangular to polar form.

On the TI-84: go to [**ANGLE**] (or [**2nd**] function) [**APPS**]. Scroll down to 5 or "R-Pr(" and press [**Enter**] . Next, enter the rectangular coordinates and close the parenthesis. Press [**Enter**], the "r" value appears. Scroll down to 6R-P , and the polar angle appears in decimal radian form.

Also under the [**ANGLE**] menu, commands 7 and 8 allow transformation from polar form to rectangular form.

Lesson Summary

There are multiple ways to represent a complex number. Each way has their advantages when doing specific problems. Knowing how to convert from one representation to another gives the student flexibility in being able to work with many different problems that use complex numbers.

Points to Consider

Complex numbers can be written in either rectangular form or in polar form. For what type of problem is it more advantageous to write a complex number in rectangular form? In polar form?

Review Questions

1. Convert the following complex number into polar form: a. $\sqrt{3} - i$
   b. $9\sqrt{3} + 9i$
   c. 3 - $4i$
   d. $\sqrt{5} - i$
2. Convert the following complex numbers into rectangular form: a. $3(\cos 210^\circ + i \sin 210^\circ)$
   b. $2\left(\cos \frac{\pi}{18} + i \sin \frac{\pi}{18}\right)$
   c. 4 cis $\left(\frac{7\pi}{4}\right)$
3. Convert the rectangular/polar coordinate into polar/rectangular coordinate: a. $(3, 3\sqrt{3})$
   b. (-2, 2)
   c. $\left(4, \frac{2\pi}{3}\right)$
   d. $\left(-1, \frac{5\pi}{6}\right)$

Review Answers

1. a. $2(\cos 330^\circ + i \sin 330^\circ)$b. $18(\cos 30^\circ + i \sin 30^\circ)$
   c. $5(\cos 306.9^\circ + \text{i} \sin 306.9^\circ)$
   d. $\sqrt{6}(\cos 335.9^\circ + i \sin 335.9^\circ)$
2. a. $\frac{-3\sqrt{3}}{2} - \frac{3}{2}i$b. $1.97 + 0.35i$
   c. $2\sqrt{2} - 2\sqrt{2}i$
3. .

5.7 Product and Quotient Theorems

Learning Objectives

- Use the rcis form of complex numbers to multiply two complex numbers
- Use the rcis form of complex numbers to divide two complex numbers
- Use the rectangular form and technological tools to check the results of multiplication and division with complex numbers

Product Theorem

Since complex numbers can be transformed in polar form, the multiplication of complex numbers can also be done in polar form. Suppose we know $z_1 = r_1\ (\cos\ _1 + i \sin\ _1)\ z_2 = r_2\ (\cos\ _2 + i \sin\ _2)$

To multiply the two complex numbers in polar form:

$z_1 \cdot z_2 = r_1(\cos\ \theta_1 + i \sin\ \theta_1) \cdot r_2(\cos\ \theta_2 + i \sin\ \theta_2)$

$= r_1 r_2(\cos\ \theta_1 + i \sin\ \theta_1)(\cos\ \theta_2 + i \sin\ \theta_2)$

$= r_1 r_2 \cdot (\cos\ \theta_1\ \cos\ \theta_2 + i \cos\ \theta_1\ \sin\ \theta_2 + i \sin\ \theta_1\ \cos\ \theta_2 + i^2 \sin\ \theta_1\ \sin\ \theta_2)$

$= r_1 r_2(\cos\ \theta_1\ \cos\ \theta_2\ + i \cos\ \theta_1\ \sin\ \theta_2 + i \sin\ \theta_1\ \cos\ \theta_2\ - \sin\ \theta_1\ \sin\ \theta_2)$

$= r_1 r_2\ (\cos\ \theta_1\ \cos\ \theta_2 - \sin\ \theta_1\ \sin\ \theta_2 + i \cos\ \theta_1\ \sin\ \theta_2 + i \sin\ \theta_1\ \cos\ \theta_2)$

$= r_1 r_2\ ([\cos\ \theta_1\ \cos\ \theta_2 - \sin\ \theta_1\ \sin\ \theta_2] + i[\cos\ \theta_1\ \sin\ \theta_2 + \sin\ \theta_1\ \cos\ \theta_2])$

(Use i^2 = -1 Gather like terms Factor out i Substitute the angle sum formulas for both sine and cosine)

$z_1 \cdot z_2 = r_1 r_2\ \text{cis}\ [(\theta_1 + \theta_2)]$

$z_1 \cdot z_2 = r_1 r_2\ \text{cis}\ [(\theta_1 + \theta_2)]$

In this last equation the product of two complex numbers in polar form can be obtained by multiplying the polar r values of each of the complex numbers and then multiplying that value by cis of the sum of each of the two angles of the individual complex numbers. This appears to be simpler than the rectangular form for multiplication of complex numbers.

Quotient Theorem

Dividing complex numbers in polar form can be shown using a similar proof that was used to show multiplication of complex numbers. Here we omit the proof and give the result. For $z_1 = r_1(\cos\ _1 + i \sin\ _1)$ and $z_2 = r_2(\cos\ _2 + i \sin\ _2)$, then $\frac{z_1}{z_2} = \frac{r_1}{r_2} \times \text{cis}\ [\theta_1 - \theta_2]$

Using the Product and Quotient Theorem

Example: Multiply $z_1 \cdot z_2$ where $z_1 = 2 + 2i$ and $z_2 = 1 - \sqrt{3}i$

For z_1,

$r_1 = \sqrt{2^2 + 2^2}$

$= \sqrt{8}$

$= 2\sqrt{2}$

and

$\tan\ \theta_1 = \frac{2}{2}$

$\tan\ \theta_1 = 1$

$\theta_1 = \frac{\pi}{4}$

Note that $\ _1$ is in the first quadrant since a, and $b > 0$.

For z_2,

$r_2 = \sqrt{1^2 + (-\sqrt{3})^2}$

$= \sqrt{1 + 3}$

$= \sqrt{4}$

$= 2$

and

$\tan \theta_2 = \frac{-\sqrt{3}}{1}$,

$\theta_2 = \frac{5\pi}{3}$

Now we can use the formula, $z_1 \quad z_2 = r_1 r_2 \text{ cis } [(\ _1 + \ _1)]$. Substituting

$z_1 \cdot z_2 = 2\sqrt{2} \times 2 \text{ cis } \left[\frac{\pi}{4} + \frac{5\pi}{3}\right]$

$= 4\sqrt{2} \text{ cis } \left[\frac{23\pi}{12}\right]$

So we have

$z_1 \cdot z_2 = 4\sqrt{2}\left(\cos \frac{23\pi}{12} + i \sin \frac{23\pi}{12}\right)$

Re-writing in approximate decimal form:

5.656 (0.966 – 0.259i)

5.46 - 1.46i

If the problem was done using only rectangular units then

$z_1 \times z_2 = (2 + 2i)(1 - \sqrt{3}i)$ or

$= 2 - 2\sqrt{3}i + 2i - 2\sqrt{3}i^2$

Gathering like terms and using $i^2 = -1$

$= (2 + 2\sqrt{3}) - (2\sqrt{3} + 2)i$

or

$5.46 - 1.46i$

Example: Using polar multiplication, find the product $(6 - 2\sqrt{3}i)(4 + 4\sqrt{3}i)$

Let $z_1 = 6 - 2\sqrt{3}i$ and $z_2 = 4 + 4\sqrt{3}i$

$r_1 = \sqrt{(6)^2 - (2\sqrt{3})^2}$ and $r_2 = \sqrt{(4)^2 + (4\sqrt{3})^2}$

$r_1 = \sqrt{36 + 12} = \sqrt{48} = 4\sqrt{3}$ and $r_2 = \sqrt{16 + 48} = \sqrt{64} = 8$

For $\ _1$, first find $\tan \theta_{ref} = \left|\frac{y}{x}\right|$

$\tan \theta_{ref} = \frac{(2\sqrt{3})}{6}$

$\tan \theta_{ref} = \frac{\sqrt{3}}{3}$

$\theta_{ref} = \frac{\pi}{6}$.

Since $x > 0$ and $y < 0$ we know that $\ _1$ is in the in the 4th quadrant:

$\theta_1 = \frac{11\pi}{6}$

For $\ _2$,

$\tan \theta_{ref} = \frac{(4\sqrt{3})}{4}$

$\tan \theta_{ref} = \sqrt{3}$,

$\theta_{ref} = \frac{\pi}{3}$

Since $\ _2$ is in the first quadrant,

$\theta_2 = \frac{\pi}{3}$

Using polar multiplication,

$z_1 \times z_2 = 4\sqrt{3} \times 8\left(\text{cis } \left[\frac{11\pi}{6} + \frac{\pi}{3}\right]\right)$

$z_1 \times z_2 = 32\sqrt{3}\left(\text{cis } \left[\frac{13\pi}{6}\right]\right)$

subtracting 2 from the augment:

$z_1 \times z_2 = 32\sqrt{3}\left(\text{cis } \left[\frac{\pi}{6}\right]\right)$

or in expanded form: $32\sqrt{3}\left(\cos\left[\frac{\pi}{6}\right] + i \sin\left[\frac{\pi}{6}\right]\right)$

In decimal form this becomes: 55.426(0.866 + 0.500i) or 48 + 27.713i

Check:

$(6 - 2\sqrt{3}i)(4 + 4\sqrt{3}i) = 24 + 24\sqrt{3}i - 8\sqrt{3}i - 24i^2$

$= 24 + 16\sqrt{3}i + 24$

$= 48 + 27.713i$

Example: Using polar division find the quotient of $\frac{z_1}{z_2}$, given that

$z_1 = 5 - 5i$

$z_2 = -2\sqrt{3} - 2i$

for $z_1 : r_1 = r_1 = \sqrt{5^2 + (-5)^2}$ or $5\sqrt{2}$ and $\tan\theta_1 = \frac{-5}{5}$, so $\theta_1 = \frac{7\pi}{4}$ (4th quadrant)

for $z_2 : r_2 = \sqrt{(-2\sqrt{3})^2 + (-2)^2}$ or $\sqrt{16} = 4$ and $\tan\theta_2 = \frac{-2}{(-2\sqrt{3})}$, so $\theta_2 = \frac{7\pi}{6}$ (3^{rd} quadrant)

Using the formula, $\frac{z_1}{z_2} = \frac{r_1}{r_2} \times \text{cis}\,[\theta_1 - \theta_2]$ or

$= \frac{5\sqrt{2}}{4} \times \text{cis } \left[\frac{7\pi}{4} - \frac{7\pi}{6}\right]$

$= \frac{5\sqrt{2}}{4} \times \text{cis } \left[\frac{7\pi}{12}\right]$

$= \frac{5\sqrt{2}}{4}\left[\cos\frac{7\pi}{12} + i \sin\frac{7\pi}{12}\right]$

$= 1.768[-0.259 + (0.966)i]$

$= -0.458 + 1.708i$

Check by using the complex conjugate to do the division in rectangular form:

$\frac{5-5i}{-2\sqrt{3}-2i} \cdot \frac{-2\sqrt{3}+2i}{-2\sqrt{3}+2i} = \frac{-10\sqrt{3}+10i+10\sqrt{3}i-10i^2}{(-2\sqrt{3})^2-(2i)^2}$

$= \frac{-10\sqrt{3}+10i+10\sqrt{3}+10}{12+4}$

$= \frac{(-10\sqrt{3}+10)+(10+10\sqrt{3})}{16}$

$= \frac{(-17.3+10)+(10+17.3)i}{16}$

$= \frac{(-7.3)+(27.3)i}{16}$ or

$-0.456 + 1.706i$

The two radically different approaches yield the same answer. The small difference between the two answers is a result of decimal rounding error.

Example: Find the quotient: $\frac{1+2i}{2-i}$

First finding the quotient by polar multiplication:

$r_1 = \sqrt{(1)^2 + (2)^2} = \sqrt{5}$ $\quad$ $r_2 = \sqrt{(2)^2 + (-1)^2} = \sqrt{5}$

$\tan \theta_1 = \frac{2}{1}$

$\tan \theta_1 = 2$

$\theta_{ref} = 1.107$ radians

since the angle is in the 1st quadrant

$_1 = 1.107$ radians

for θ_2,

$\tan \theta_2 = \frac{-1}{2}$

$\tan \theta_{ref} = \frac{1}{2}$

$\theta_{ref} = 0.464$ radians

since $_2$ is in the 4th quadrant, between 4.712 $\left(\text{or } \frac{3\pi}{2}\right)$ and 6.282 radians (or 2)

$_2 = 5.820$ *radians*

Finally, using the division formula,

$\frac{z_1}{z_2} = \frac{\sqrt{5}}{\sqrt{5}}[\text{cis } (1.107 - 5.820)]$

$\frac{z_1}{z_2} = [\text{cis } (-4.713)]$

$\frac{z_1}{z_2} = [\cos (-4.713) + i \sin (-4.713)]$

$\frac{z_1}{z_2} = [\cos (1.570) + i \sin (1.570)]$

If we assume that $\frac{\pi}{2} = 1.570$, then

$\approx \frac{z_1}{z_2} = \left[\cos \left(\frac{\pi}{2}\right) + i \sin \left(\frac{\pi}{2}\right)\right]$

$\frac{z_1}{z_2} = 0 + 1i = 1$

To check, multiply numerator and denominator by the denominator's complex conjugate:

$\frac{1+2i}{2-i} \cdot \frac{2+i}{2+i} = \frac{2+i+4i+2i^2}{4+2i-2i-i^2}$

$= \frac{2+5i-2}{4+1}$

$= \frac{5i}{5}$

$= i$

Applications, Technological Tools

After utilizing product or quotient laws, answers can be checked using a graphing calculator. Convert the given values into $a + bi$ form and then put the calculator into "pol" mode. Perform the operation.

Lesson Summary

There are two distinct methods when multiplying or dividing two complex numbers- one utilizes the complex number written in a + bi form, while the other uses the polar form. However, there is really only one practical method for finding a power or a root of a complex number and that is utilizing the polar form, or cis form, for the complex number.

Points to Consider

Some students may be tempted to ask why is this new method presented- it is longer and (in this example) messier when there is an easier, less messier way of doing it? To give encouragement to learn alternative methods for solving problems, let's look at science history, in particular Einstein and the Theory of Special Relativity.

Prior to 1905 when Special Relativity was published, motion was understood by Newton's Laws of Motion and Galileo's kinematics equations. For over 300 years whenever scientists used these laws to explain phenomena that involved motion their results were just about perfect. Yet by the beginning of the 20^{th} Century, Einstein glimpsed at some discrepancies within the laws of motion and the laws of electromagnetism and realized something might be wrong with the way physicists were looking at the world. He developed a method that would bring these powerful laws together without discrepancies. What came out of his efforts was the Theory of Special Relativity. A cynic might have asked at that time, what was so important about learning a newer, more complicated method, when the older and reliable methods worked just as well? The answer for understanding Relativity was, yes, the new method yielded the identical answers as the older method, BUT the newer method was able to answer modern questions, not asked of the older methods. In other words it could do all that the older methods can do but also more. This idea has several examples in science and mathematics and is known as the *Correspondence Principle.*

This principle is true of the method of multiplying complex numbers. As seen, using the newer method yields identical results as the older method. After becoming proficient in using the newer method, problems will be presented that when using this newer method, will take much less time then the older method took, that is if the older method could even do the problem!

Review Questions

1. a. Find the product using polar form: $(2 + 2i)(\sqrt{3} - i)$b. Multiply: $2(\cos 40^\circ + i \sin 40^\circ) \times 4(\cos 20^\circ + i \sin 20^\circ)$
   c. Multiply: $2\left(\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}\right) \times 2\left(\cos \frac{\pi}{10} + i \sin \frac{\pi}{10}\right)$
   d. Divide: $2(\cos 80^\circ + i \sin 80^\circ) \div 6(\cos 200^\circ + i \sin 200^\circ)$
   e. Divide: $3\text{cis}(130^\circ) \div 4\text{cis}(270^\circ)$

Review Answers

1. a. $4\sqrt{2}(\cos 15^\circ + i \sin 15^\circ)$b. 8 cis (60°)
   c. 4 cis $\left(\frac{9\pi}{40}\right)$
   d. $\frac{1}{3}\text{cis}\ (240^\circ)$
   e. $\frac{3}{4}\text{cis}\ (220^\circ)$

5.8 Powers and Roots of Complex Numbers

Learning Objectives

- Use De Moivre's Equation to find powers of complex numbers
- Prove De Moivre's Theorem
- Use De Moivre's theorem to find all solutions to an equation

Introduction

Whether it is adding, subtracting, multiplying, dividing or some other mathematical operation that is being done on two or more complex numbers, there will be more than one method- using rectangular form or polar form

De Moivre's Theorem

How do we raise a complex number to a power? Let's start with an example

$(-4 - 4i)^3 = (-4 - 4i)$ x $(-4 - 4i)$ x $(-4 - 4i) =$

In rectangular form, this can get very complex. What about in r cis form?

$(-4 - 4i) = 4\sqrt{2}\ cis\left(\frac{5\pi}{4}\right)$

So the problem becomes

$4\sqrt{2}\ \text{cis}\left(\frac{5\pi}{4}\right) \cdot\ 4\sqrt{2}\ \text{cis}\left(\frac{5\pi}{4}\right) \cdot\ 4\sqrt{2}\ \text{cis}\left(\frac{5\pi}{4}\right)$

and using our multiplication rule from the previous section,

$(-4 - 4i)^3 = (4\sqrt{2})^3\ \text{cis}\left(\frac{15\pi}{4}\right)$

Notice, $(a + bi)^3 = r^3\ cis\ 3$

In words: Raise the r-value to the same degree as the complex number is raised and then multiply that by *cis* of the angle multiplied by the number of the degree.

Based on the example above, it appears that we can generalize

$(a + bi)^n = r\ cis(n) = r(\cos(n) + i\ \sin(n))$

If this is correct, then the polar form provides a much faster result for raising a complex number to a power than doing the problem in rectangular form.

Example: Find the value of $(1 + \sqrt{3}i)^4$

$r = \sqrt{(1)^2 + (\sqrt{3})^2} = 2$

$\tan\theta_{ref} = \frac{\sqrt{3}}{1}$,

and is in the 1st quadrant, so

$\theta = \frac{\pi}{3}$

Using our equation from above:

$z^4 = r^4\ cis\ 4\theta$

$z^4 = (2)^4\ cis\ 4\frac{\pi}{3}$

Expanding cis form:

$z^4 = 16\left(\cos\left(\frac{4\pi}{3}\right) + i\ \sin\left(\frac{4\pi}{3}\right)\right)$

$= 16\left((-0.5) - 0.866i\right)$

Finally we have

$z^4 = -8 - 13.856i$

Check by complex rectangular multiplication. First we will find z^2, then square that result:

$(1 + \sqrt{3}i)^2 = 1 + 2\sqrt{3}i - 3$

$= -2 + 2\sqrt{3}i$

Now by squaring the square, the fourth power is found or $(-2 + 2\sqrt{3}i)^2$

$4 - 8\sqrt{3}i - 12 = -8 - 8\sqrt{3}i$

$= -8 - 13.856i$

The general rule for raising a complex number to any power is stated by **De Moivre's**

Equation:

Table 5.14:

Let $z = r(\cos\ \ + i \sin\ )$ be a complex number in rcis form. If n is a positive integer, z^n is $z^n = r^n\ (\cos(n\ ) + i \sin(n\ ))$

Proof:

The proof of De Moivre's equation uses **mathematical induction**. Mathematical induction is a method of proving something is true for an infinite number of cases. In this example, we want to show that De Moivre's equation works for all integer values of n.

The way we do this (and any induction proof) is first we will check that the theorem is true for $n = 1$. Then we assume it is true for $n = k$ (called the induction step). Finally, we use the case $n = k$ to show that if the theorem is true for $n = k$ then it is also true for $n = k + 1$. Once we have done this, we have proven the theorem always works for any integer n.

If n = 1, then the theorem is

$z^1 = r^1\ (\cos(1 \cdot\ ) + i \sin(1 \cdot\ ))$

Which is the rcis form of z

Now we assume that

$z^k = r^k\ (\cos(k\ ) + i \sin(k\ ))$

Then we check the case of $z^{k\,+\,1}$.

$z^{k\,+\,1} = z^k\ z$

By the induction step,

$z^{k+1} = r^k(\cos(k\theta) + i \sin(k\theta))r(\cos(\theta) + i \sin(\theta))$

$= r^{k+1}(\cos(k\theta) + i \sin(k\theta))(\cos(\theta) + i \sin(\theta))$

$= r^{k+1}(\cos(k\theta)\cos(\theta) + i \sin(k\theta)\cos(\theta) + i \cos(k\theta)\sin(\theta) + i^2 \sin(k\theta)\sin(\theta))$

Gathering like terms and using i^2 = -1

$= r^{k+1}(\cos(k\theta)\cos(\theta) - \sin(k\theta)\sin(\theta) + i\ (\sin(k\theta)\cos(\theta) + \cos(k\theta)\sin(\theta)))$

Finally, we use the angle addition identities, and we get

$= r^{k+1}(\cos(k\theta + \theta) + i \sin(k\theta + \theta))$

$= r^{k+1}(\cos((k + 1)\theta) + i \sin((k + 1)\theta))$

And now we are done. We have shown for any positive integer n,

$z^n = r^n(\cos(n\theta) + i \sin(n\theta))$.

nth Root Theorem

Often in mathematics whenever an operation is presented, the inverse operation follows. This will be the case for the current topic. The inverse operation of finding a power for a number is to find a root of the same number.

The problem, for example is find the square root of $(1 + i)$.

Previously you have learned that the square root of a number, x, was shown to equal $x^{1/2}$.

Any root, say the nth root can be written as $x^{1/n}$

The formula for De Moivre's theorem also works for fractional powers (though we won't prove that here). That means that this same formula can be used for finding roots:

$z^{1/n} = (a + bi)^{1/n} = r^{1/n} cis\left(\frac{\theta}{n}\right)$

Example: Find $\sqrt{1+i}$

Fist, rewriting in exponential form: $(1 + \text{i})^{½}$

And now in polar form:

$\sqrt{1+i} = \left(\sqrt{2}\ cis\left(\frac{\pi}{4}\right)\right)^{1/2}$

Expanding cis form,

$= \left(\sqrt{2}\left(\cos\left(\frac{\pi}{4}\right) + i\ \sin\left(\frac{\pi}{4}\right)\right)\right)^{1/2}$

Using the formula:

$= (2^{1/2})^{1/2}\left(\cos\left(\frac{1}{2}\cdot\frac{\pi}{4}\right) + i\ \sin\left(\frac{1}{2}\cdot\frac{\pi}{4}\right)\right)$

$= 2^{1/4}\left(\cos\left(\frac{\pi}{8}\right) + i\ \sin\left(\frac{\pi}{8}\right)\right)$

In decimal form, we get

$=1.189(\ 0.924 + 0.383i)$

$=1.099 + 0.455\text{i}$

To check, we will multiply the result by itself in rectangular form:

$(1.099 + 0.455i)\ \cdot\ (1.099 + 0.455\text{i}) = 1.099^2 + 1.099(0.455i) + 1.099(0.455i) + (0.455i)^2$

$= 1.208 + 0.500i + 0.500i + 0.208i^2$

$= 1.208 + i - 0.208$ or

$= 1 + i$

The Fundamental Theorem of Algebra states that the number of solutions for an equation is equal to the degree of an equation. Finding $\sqrt{1+i}$, is the same as solving for x in $x^2 = 1 + i$.

That would mean that there are two values for x, and in the solution above, but we only found one value. What is the other value? How is it found?

If the square root of a real number, n, is b, then both b^2 and $(-b)^2$ will equal n. Can this also be true for complex numbers?

In the above example $\sqrt{1+i}$ is equivalent to $\sqrt{\sqrt{2}}\ \left(\cos\left[\frac{\pi}{8}\right] + i\ \sin\left[\frac{\pi}{8}\right]\right)$,

or in approximate decimal form: $1.099 + 0.455i$.

The opposite of this number is $(-1.099 - 0.455i)$.

Squaring the opposite,

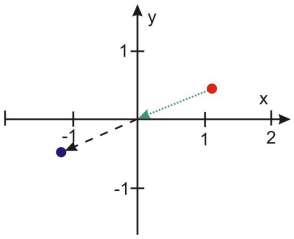
$(-1.099 - 0.455\text{i})^2 = (-1.099)^2 + (-1.099)(-0.455i) + (-1.099)(-0.455i) + (-0.455i)^2$

$= 1.208 + 0.500i + 0.500i + 0.208i^2$

$= 1.208 + i - 0.208$

$= 1 + i$

Graphing both of these complex solutions on the same polar axis reveals an important feature of the roots of complex numbers:



What this reveals is that both solutions are the same distance from the origin, and but they are rotated by radians when graphed. This suggests is that the roots of complex numbers, when graphed, are spaced apart in such a manner, that the difference between their angles is equal. In this case, both solutions are separated by radians. Finding the cube root of a complex number would have the three solutions be $\frac{2\pi}{3}$ radians or 120° apart when graphed. Finding the fourth root of a complex number would have the four solutions be $\frac{2\pi}{4} = \frac{\pi}{2}$ radians or 90° apart.

The general rule would be to find the first root by the method seen in the above example. Leave the answer in cis form. To find the additional roots use rotations of $\frac{2\pi}{n}$ radians or $\frac{360^{\circ}}{n}$ to find the angle difference between the roots. Add this angle "n - 1" times to the first angle, until a complete circle has been formed.

The procedures, when combined, are known as De Moivre's Theorm.

Solve Equations using De Moivre's Theorem

Example: Find the value of $x : x^3 = \left(1 - \sqrt{3}i\right)$

First we put $1 - \sqrt{3}i$ in polar form.

Use $x = 1,\ y = -\sqrt{3}$ to obtain $r = 2,\ \theta = \frac{5\pi}{3}$

let $z = \left(1 - \sqrt{3}i\right)$ in rectangular form

$z = 2$ cis $\left(\frac{5\pi}{3}\right)$ in polar form

$x = \left(1 - \sqrt{3}i\right)^{1/3}$

$x = \left[2cis\left(\frac{5\pi}{3}\right)\right]^{1/3}$

Use De Moivre's Equation to find the first solution:

$x_1 = 2^{1/3}cis\left(\frac{5\pi/3}{3}\right)$ or $2^{1/3}cis\left(\frac{5\pi}{9}\right)$

Leave answer in cis form to find the remaining solutions:

$n = 3$ which means that the 3 solutions are $\frac{2\pi}{3}$ radians apart or

$x_2 = 2^{1/3}cis\left(\frac{5\pi}{9} + \frac{2\pi}{3}\right)$ and $x_3 = 2^{1/3}cis\left(\frac{5\pi}{9} + \frac{2\pi}{3} + \frac{2\pi}{3}\right)$

Note: It is not necessary to add $\frac{2\pi}{3}$ again. Adding $\frac{2\pi}{3}$ three times equals 2 . That would result in rotating

around a full circle and to start where it all began- that is the first solution.

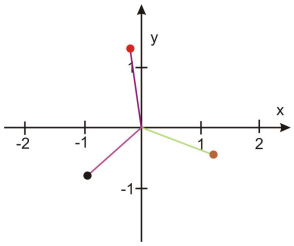
The three solutions are:

$x_1 = 2^{1/3} cis\left(\frac{5\pi}{9}\right)$

$x_2 = 2^{1/3} cis\left(\frac{11\pi}{9}\right)$

$x_3 = 2^{1/3} cis\left(\frac{17\pi}{9}\right)$

Each of these solutions, when graphed will be $\frac{2\pi}{3}$ apart.



Check any one of these solutions, for example the second solution, to see if the results are confirmed. To check the first and third solution will be given as an exercise at the back of the chapter.

$x_2 = 2^{1/3} cis\left(\frac{11\pi}{9}\right)$

$= 1.260\left[\cos\left(\frac{11\pi}{9}\right) + i\,\sin\left(\frac{11\pi}{9}\right)\right]$

$= 1.260[-0.766 - 0.643i]$

$= -0.965 - 0.810i$

Does $(-0.965 - 0.810i)^3$ or $(-0.965 - 0.810i)\ (-0.965 - 0.810i)\ (-0.965 - 0.810i)$

$= \left(1 - \sqrt{3}i\right)$?

Using polar multiplication would in effect be the reverse of De Moivre's Theorem. Using rectangular or algebraic multiplication would yield an independent check!

$(-0.965 - 0.810i)\ (-0.965 - 0.810i) = 0.931 + 0.782i + 0.782i + 0.656i^2$

or

$(0.932 - 0.656) + 2(0.782)i$

or

$0.276 + 1.564i$

Now multiply again

$(0.276 + 1.564i)(-0.965 - 0.810i) = -0.266 - 0.224i - 1.509i - 1.266i^2$ or

$(-0.266 + 1.266) - (0.224 + 1.509)$i or

$1 - 1.733i$ or approximately $1 - \sqrt{3}i$

Example: What are the two square roots of i? Let $z = \sqrt{0 + i}$

$r = 1,\ \theta = \pi/2$ or $z = \left[1 \times\ cis\frac{\pi}{2}\right]^{1/2}$ Utilizing De Moivre's Theorem:

$z_1 = \left[1 \times\ cis\frac{\pi}{4}\right]$ or $z_2 = \left[1 \times\ cis\frac{5\pi}{4}\right]$

$z_1 = 1\left(\cos\frac{\pi}{4} + i\,\sin\frac{\pi}{4}\right)$ or $z_2 = 1\left(\cos\frac{5\pi}{4} + i\,\sin\frac{5\pi}{4}\right)$

$z_1 = 0.707 + 0.707i$ or $z_2 = -0.707 - 0.707i$

Check for z_1 solution: $(0.707 + 0.707i)^2$ = i?

$0.500 + 0.500i + 0.500i + 0.500i^2 = 0.500$ + i + 0.500(-1) or i

Example: What are the four fourth roots of 1?

Let $z = 1$ or $z = 1 + 0i$ Then the problem becomes find $z^{1/4} = (1 + 0i)^{1/4}$

Since $r = 1$ $\theta = 0$, $z^{1/4} = [1 \times cis\ 0]^{1/4}$ with $z_1 = 1^{1/4}\left(\cos\ \frac{0}{4} + i\ \sin\ \frac{0}{4}\right)$ or $1(1 + 0)$ or 1

That root is not a surprise. Now use De Moivre's to find the other roots:

$z_2 = 1^{1/4}\left[\cos\left(0 + \frac{\pi}{2}\right) + i\ \sin\left(0 + \frac{\pi}{2}\right)\right]$ Since there are 4 roots, dividing 2 by 4 yields 0.5

or 0 + i or just i $z_3 = 1^{1/4}\left[\cos\left(0 + \frac{2\pi}{2}\right) + i\ \sin\left(0 + \frac{2\pi}{2}\right)\right]$ which yields z_3 = -1

Finally $z_4 = 1^{1/4}\left[\cos\left(0 + \frac{3\pi}{2}\right) + i\ \sin\left(0 + \frac{3\pi}{2}\right)\right]$ or $z_4 = -i$

The four fourth roots of 1 are 1, i, -1 and -i

Example: What are the three cube roots of 1?

Let $z = 1$ or $z = 1 + 0i$ Since r = 1 $= 0$ Then the problem becomes find $z^{1/3} = [1 \times \text{cis } 0]^{1/3}$

$z_1 = 1^{1/3}[\cos(0) + i\ \sin(0)]$ or 1

$z_2 = 1^{1/3}\left[\cos\left(0 + \frac{2\pi}{3}\right) + i\ \sin\left(0 + \frac{2\pi}{3}\right)\right]$ or $-0.500 + 0.866i$

$z_3 = 1^{1/3}\left[\cos\left(0 + \frac{4\pi}{3}\right) + i\ \sin\left(0 + \frac{4\pi}{3}\right)\right]$ or $-0.500 + 0.866i$

Check z_2 : $(-0.500 + 0.866i)^3 = 1$?

$(-0.500 + 0.866i)(-0.500 + 0.866i) = 0.25$ - 0.866i + $.750i^2$ or -0.500 - 0.866i

$(-0.500 - 0.866i)(-0.500 + 0.866i) = 0.25 - 0.866i + 0.866i$ -0.750 i^2 or 1.

Applications, Technological Tools

After utilizing De Moivre's Theorem, answers can be checked using a graphing calculator. Convert the given values into $a + bi$ form and then put the calculator into "pol" mode. Perform the operation.

Lesson Summary

There is only one practical method for finding the power or roots of a complex number. Before raising a complex number to a power or finding a root or roots of a complex number, the complex number must be in polar or cis form, and then De Moivre's Theorem can be utilized.

Points to Consider

Having a number in polar form can yield results that the rectangular form could NOT! Recall the *Correspondence Principle*. Sometimes a method learned at an earlier level in mathematics will work up to a point. When that method will no longer work, newer, sometimes more complicated, or seemingly more complicated methods are introduced, often to the protestations of students. Often these new methods are helpful for solving a wider range of problems.

Review Questions

1. Perform indicated operation on these complex numbers: a. Divide: $\frac{2+3i}{1-i}$
   b. Multiply: $(-6-i)(-6+i)$
   c. Multiply: $\left(\frac{\sqrt{3}}{2}-\frac{1}{2}i\right)^2$
   d. Find the product using polar form: $(2+2i)(\sqrt{3}-i)$
   e. Multiply: $2(\cos 40^\circ + i \sin 40^\circ) \times 4(\cos 20^\circ + i \sin 20^\circ)$
   f. Multiply: $2\left(\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}\right) \times 2\left(\cos \frac{\pi}{10} + i \sin \frac{\pi}{10}\right)$
   g. Divide: $2(\cos 80^\circ + i \sin 80^\circ) \div 6(\cos 200^\circ + i \sin 200^\circ)$
   h. Divide: $3 \text{ cis}(130^\circ) \div 4 \text{ cis}(270^\circ)$
2. Use De Moivre's Theorem: a. $[3(\cos 80^\circ + i \sin 80^\circ)]^3$
   b. $\left[\sqrt{2}\left(\cos \frac{5\pi}{16} + i \sin \frac{5\pi}{16}\right)\right]^4$
   c. $\left(\sqrt{3}-i\right)^6$
   d. The 3 complex cube roots of $1 + i$
   e. The 4 complex fourth roots of - $16i$
   f. The five complex fifth roots of i

Review Answers

1. a. $\frac{-1}{2}+\frac{5}{2}i$ b. 37
   c. $\frac{1}{2}+\frac{-\sqrt{3}}{2}i$
   d. $4\sqrt{2}(\cos 15^\circ + i \sin 15^\circ)$
   e. $8 \text{ cis}(60^\circ)$
   f. $4 \text{ cis}\left(\frac{9\pi}{40}\right)$
   g. $\frac{1}{3} \text{ cis}(240^\circ)$
   h. $\frac{3}{4} \text{ cis}(220^\circ)$
2. a. $-\frac{27}{2}-\frac{27\sqrt{3}}{2}i$ b. $-2\sqrt{2}-2\sqrt{2}i$
   c. -64
   d. $\left(\sqrt[6]{2}\right)\text{cis } 15^\circ$, $\sqrt[6]{2} \text{ cis } 135^\circ$, $\sqrt[6]{2} \text{ cis } 255^\circ$,
   e. 2cis 67.5º, 2cis 157.5º, 2cis 247.5º, 2cis 337.5º
   f. cis 18º, cis 90º, cis 162º, cis 234º, cis 306º

Chapter 6

Vectors

6.1 Vectors in a Plane

Learning objectives

- Identify difference between vector and scalar quantities.
- Represent vector quantities graphically and using component notation.
- Compute the resultant of a set of vectors.
- Determine the unit vector to represent a given direction.

Introduction

In many cases, numbers alone are sufficient to model physical situations or to answer such questions as, "what temperature is it?" or "how much lumber will I need to fix the porch stairs?" or "what grade should I expect in this class?" Sometimes, however, a single number is not sufficient to convey the necessary information. For example, crime scene investigators know that both the *speed* and the *direction* of a car's motion will affect the outcome of a crash. Quantities that can be represented by a single number are called **scalar** quantities. Those that require multiple numbers or directional representation are called **vectors.** In this section we will discuss both the geometric and algebraic properties of vectors in two dimensions.

Two Dimensional Vectors

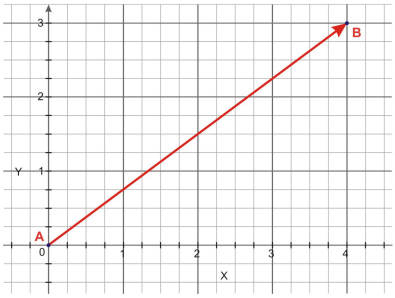
Throughout this website we will be using bold-face type to indicate vector quantities and regular face type for scalar quantities, whether we are using the endpoints of the vector or another variable to represent the vector quantity: $\boldsymbol{AB}$ or $\boldsymbol{r}$. Another way to symbolize vector quantities is to use letters with arrows over them. For example, physicists use $\overrightarrow{v}$ to indicate the velocity of an object.

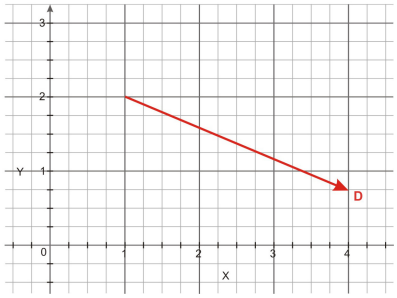
Vectors can be represented graphically using an arrow which points from the vector's initial point to its terminal point. The length of the arrow-shaft represents the **magnitude** of the vector. The arrowhead identifies the vector's direction.

Vectors $\boldsymbol{AB}$ and $\boldsymbol{CD}$ have the same magnitude and the same direction, so we can say that $\boldsymbol{AB} = \boldsymbol{CD}$,

even if they don't have the same location in space. Vectors $\boldsymbol{AB}$ and $\boldsymbol{EF}$ have the same magnitude, but they have opposite directions, therefore $\boldsymbol{EF} = -\boldsymbol{AB}$. The magnitude of a vector is a scalar quantity which can be symbolized by $|\ \boldsymbol{AB}\ |$, $|\ r\ |$, $|\vec{v}|$, or v.

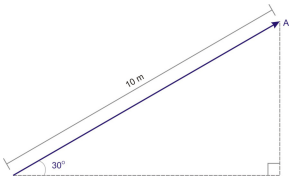
Vectors can also be represented algebraically, using **component notation.** The components tell us the extent of the vector along the coordinate directions. For example, the vector on the left in the image below has components in both the x-direction, $\boldsymbol{AB}_x$, and the y-direction, $\boldsymbol{AB}_y$. Here, $\boldsymbol{AB}_{\mathrm{x}} = 4$ and $\boldsymbol{AB}_y = 3$ because point B is 3 units to the right and 3 units up from point A. Similarly, vector $\boldsymbol{D}$ has components $\boldsymbol{D}_x = 3$ and $\boldsymbol{D}_y = -1.25$. The negative sign in $\boldsymbol{D}_y$ indicates that the y-component of vector $\boldsymbol{D}$ is downward, in the negative y direction. As with other numbers, we usually only include the negative sign explicitly.





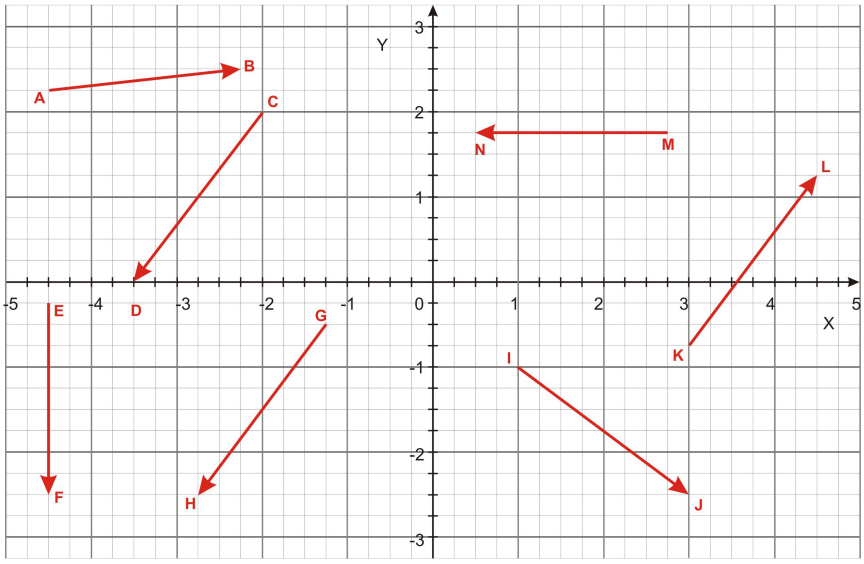
Sometimes bracket notation is used to identify the components of a vector. The components of the vector are written as an ordered set of values, similar to the coordinates of a point in space. For example, vector $AB = \langle 4, 3 \rangle$ and $D = \langle 3, -1.25 \rangle$.

If the coordinate system is polar rather than rectangular, a vector is identified by its magnitude and its direction with respect to the x-axis (or r-axis). In the case of the vector below, $\boldsymbol{A}$ = 10 m @ 30°.



The process of determining the components of a vector is also known as **resolving** the vector. For example, if we wish to convert the radial coordinate notation for vector $\boldsymbol{A}$ into rectangular coordinates, we can use right triangle trigonometry to resolve the vector. First, create a right triangle with the vector as the hypotenuse and with the two legs of the triangle parallel to the rectangular coordinate axes. Then, according to the definitions of sine and cosine, $\boldsymbol{A}_x = |\ \mathrm{A}\ | \cos 30$ and $\boldsymbol{A}_y = |\ \mathrm{A}\ | \sin 30$.

Example: What are the x and y components of the vectors shown in the diagram?



Solution:

Table 6.1:

	AB	***CD***	***EF***	***GH***	***IJ***	***KL***	***MN***
***x* - component**	2.25	-1.5	0	-1.5	2.0	1.5	-2.25
***y* - component**	0.25	-2.0	-2.25	-2.0	-1.5	2.0	0

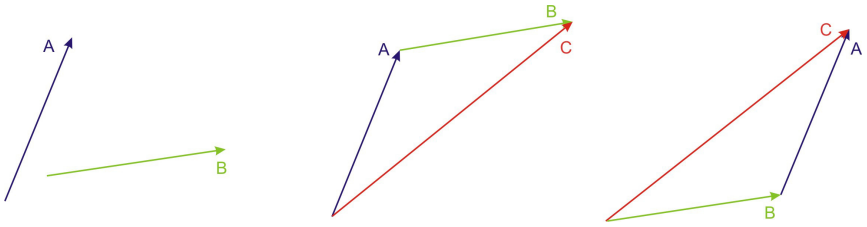
In the diagram, each division is 0.25 units. All vectors which point toward the left have negative x-components and those that point downward have negative y-components. Notice that for the horizontal vector, $\boldsymbol{MN}$, the y-component is equal to 0. Likewise, for the vertical vector, $\boldsymbol{EF}$, the x-component is equal to 0.

Example: (a) Which of the vectors in Example #1 is equal to vector $\boldsymbol{CD}$? (b) Which vector is equal to $-\boldsymbol{CD}$?

Solution: (a) Vector $\boldsymbol{GH} = \boldsymbol{CD}$. Both vectors have the same length and the same orientation. (b) Vector $\boldsymbol{KL} = -\boldsymbol{CD}$. Both vectors have the same length and the two vectors point in opposite directions.

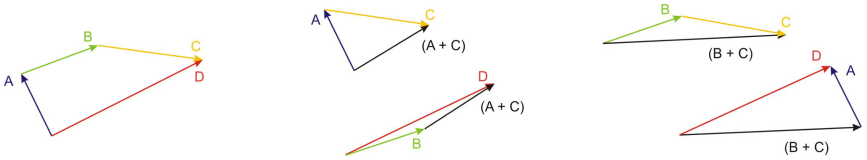
Addition and Subtraction of Vector Quantities

Vectors, like scalars, can be combined arithmetically, but, in contrast to adding scalars, we need to take the direction of vectors into account. In the diagram below, vector $\boldsymbol{C}$ is the **resultant** or **sum** of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$. $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C}$. To add the vectors graphically, position the individual vectors head to tail without changing their orientations. By head-to-tail, we mean that the initial point of the second vector is coincident with the terminal point of the first vector and so on. As you can see in the figure below, the order of the head-to-tail addition of the vectors does not affect the magnitude or direction of the resultant. In other words, the addition of the vectors is commutative: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$.

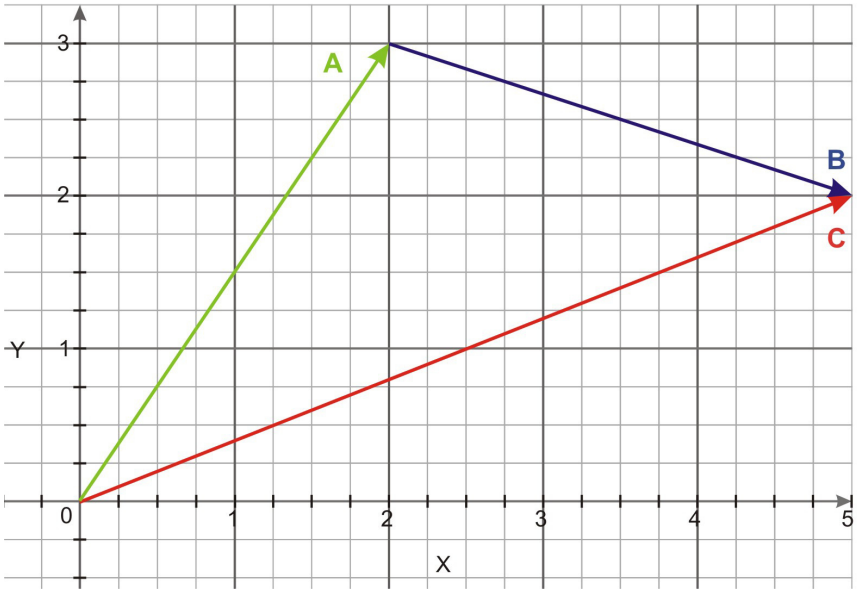


The head-to-tail addition process also works to combine groups of three or more vectors. In the diagram below, we can also see that the associative law also applies to vector addition:

$\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = (\boldsymbol{A} + \boldsymbol{C}) + \boldsymbol{B} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$.



Vectors can be added graphically, as in the diagrams above. They can also be added using their components. For any set of vectors, the x-component of the resultant vector is equal to the sum of the x-components of the individual vectors. Likewise, the y-component of the resultant is the sum of the individual y-components. For example, in the diagram below, $\overrightarrow{A} + \overrightarrow{B} = \overrightarrow{C}$. In component notation, $\overrightarrow{A} = \langle 2, 3 \rangle$ and $\overrightarrow{B} = \langle 3, -1 \rangle$. To determine the components of the resultant vector, we sum the x and y components: $\overrightarrow{C} = \langle (2 + 3), (3 - 1) \rangle = \langle 4, 2 \rangle$



Example: Demonstrate the associative law using vectors $\overrightarrow{AB}$, $\overrightarrow{CD}$, and $\overrightarrow{DE}$ from Example 1. Remember that $\overrightarrow{AB} = \langle 2.25, 0.25 \rangle$, $\overrightarrow{CD} = \langle -1.5, -2 \rangle$, and $\overrightarrow{EF} = \langle 0, -2.25 \rangle$.

Solution: The associative law states that $\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = (\boldsymbol{A} + \boldsymbol{C}) + \boldsymbol{B} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$. Here, that means that $(\overrightarrow{AB} + \overrightarrow{CD}) + \overrightarrow{EF} = \overrightarrow{AB} + (\overrightarrow{CD} + \overrightarrow{EF}) = \overrightarrow{AB} + \overrightarrow{CD} + \overrightarrow{EF}$.

Firstly, $(\overrightarrow{AB} + \overrightarrow{CD}) = \langle (2.25 - 1.5), (0.25 - 2) \rangle = \langle 0.75, -1.75 \rangle$

then, $(\overrightarrow{AB} + \overrightarrow{CD}) + \overrightarrow{EF} = \langle (0.75 + 0), (-1.75 - 2.25) \rangle = \langle 0.75, -4 \rangle$

Compare this to the other order for the addition.

Firstly, $(\overrightarrow{CD} + \overrightarrow{EF}) = \langle(-1.5 + 0), (-2 - 2.25)\rangle = \langle -1.5, -4.25\rangle$

then, $\overrightarrow{AB} + (\overrightarrow{CD} + \overrightarrow{EF}) = \langle(2.25 - 1.5), (0.25 - 4.25)\rangle = \langle 0.75, -4\rangle$

Also, $\overrightarrow{AB} + \overrightarrow{CD} + \overrightarrow{EF} = \langle(2.25 - 1.5 + 0), (0.25 - 2 - 2.25)\rangle = \langle 0.75, -4\rangle$

Example

My aunt Frances and her son Allen both have homes in the hamlet of Blue Lake. Usually, my cousin drives around the lake to reach his mother's home, but when the weather is fine, he sometimes takes his canoe across the lake instead. The black path (*A, B, C, E, D, F*) marks his driving route and the red path marks his trip in the canoe. Use vector addition techniques to determine both (a) the distance he travels across the lake and (b) his heading, . A map of the area around Blue Lake and the distances for each leg of his journey are shown in the figure below.

Solution:

Both the canoe trip and the drive will take Allen from his house to his mother's home. The vector equation $\overrightarrow{AF} = \overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD} + \overrightarrow{DE} + \overrightarrow{EF}$ describes this situation. Since the roads around Blue Lake are aligned with the compass directions, we can define our coordinate system such that $\boldsymbol{y}$ = north and $\boldsymbol{x}$ = east and put the origin of our coordinate system at point A, Allen's house. We can determine the components of the canoe trip if we know the components of each individual legs of the drive.

$AF_x = AB_x + BC_x + CD_x + DE_x + EF_x = 0 + 295 \text{ m} + 0 - 225 \text{ m} + 0 = 70 \text{ m}$

$AF_y = AB_y + BC_y + CD_y + DE_y + EF_y = -50 \text{ m} + 0 + 310 \text{ m} + 0 - 60 \text{ m} = 200 \text{ m}$

Now that we know the x and y components of vector $\boldsymbol{AF}$ we can determine the magnitude and direction of the vector. According to the Pythagorean Theorem,

$AF^2 = AF_x^2 + AF_y^2$

$AF = \sqrt{(AF_x)^2 + (AF_y)^2} = \sqrt{70^2 + 200^2} = \sqrt{4900 + 40,000} = 212\text{m}$

Then, since we know two sides of a right triangle, we can use trigonometry to determine Allen's heading.

$\tan\ \theta = \frac{AF_y}{AF_x} = \frac{200}{70} = 0.2857$ so $\quad = 70.7^\circ$ N of E

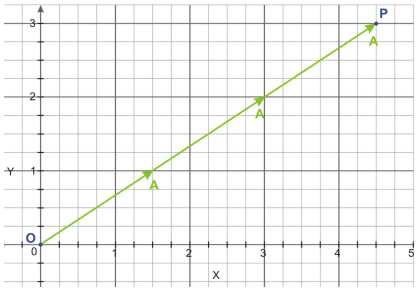
To subtract one vector from another, we simply add its opposite. For example, $\overrightarrow{A} - \overrightarrow{B} = \overrightarrow{A} + (-\overrightarrow{B})$. Compare the two sets of vectors shown below. $\overrightarrow{A} = \langle 3, 2\rangle$ and $\overrightarrow{B} = \langle 2, -1\rangle$. Therefore $\overrightarrow{A + B} = \langle(3 + 2), (2 - 1)\rangle = \langle 5, 1\rangle$, but $\overrightarrow{A - B} = \langle(3 - 2), (2 - (-1)\rangle = \langle 1, 3\rangle$.

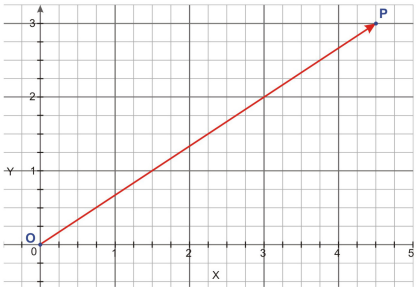
Multiplying By a Scalar

If we need to add several vectors that have the same magnitude and direction, we can utilize the fact that vectors can use a simplification commonly used when adding scalars. Specifically, multiply the scalar by the number of times it appears as a factor in the addition. For example,

$a + b + c + b + b + a + b = 2a + 4b + c$

When multiplying a vector by a positive scalar, one changes the magnitude of the vector without changing its direction. Let's look at the set of vectors shown below. Three identical vectors, **A**, placed end to end will reach from the origin to point P. This same displacement can be represented by the single vector **B**. Therefore $3\mathbf{A} = \mathbf{B}$. Here $\overrightarrow{A} = \langle 1.5, 1\rangle$ and $\overrightarrow{B} = \langle 4.5, 3\rangle$.





When a vector is multiplied by a scalar, each of the components of the vector are multiplied by that same scalar. In component notation this becomes,

$\vec{A} + \vec{A} + \vec{A} = \langle (1.5 + 1.5 + 1.5), (1 + 1 + 1) \rangle = \langle 4.5, 3 \rangle$

$3(\vec{A}) = \langle (3 * 1.5), (3 * 1) \rangle = \langle 4.5, 3 \rangle$

Vectors can be divided by scalar quantities as well, since by definition $\frac{x}{3} = \left(\frac{1}{3}\right) x$.

Example: A force is a vector quantity indicting the magnitude and direction of a push or pull on an object. Stickman Beauford and his friends Akiko and Carlos, are trying to push their car out of the mud. Beauford and Carlos push with a force of 150 N, but Akiko has a sore wrist and can only push with a force of 85 N. (Note, 1.0 N = 0.225 lbs). If all three of them push horizontally on back of the trunk, what is their total force on the car?

$\overrightarrow{F_{total}} = \overrightarrow{F_{Akiko}} + \overrightarrow{F_{Beauford}} + \overrightarrow{F_{Carlos}}$

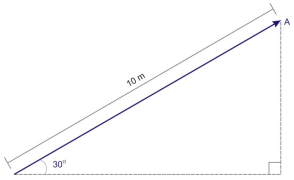
Since Beauford and Carlos push with equal force, $\overrightarrow{F_{Beauford}} + \overrightarrow{F_{Carlos}} = 2\overrightarrow{F_{Beauford}}$, therefore $\overrightarrow{F_{total}} = \overrightarrow{F_{Akiko}} + 2\overrightarrow{F_{Beauford}} = [85N + (2 * 150N)]\hat{x} = (385N)\hat{x}$

Unit Vectors

Dividing any vector by its own magnitude produces a new vector with the same direction but with unit magnitude (i.e. magnitude equal to 1), called a **unit vector.**

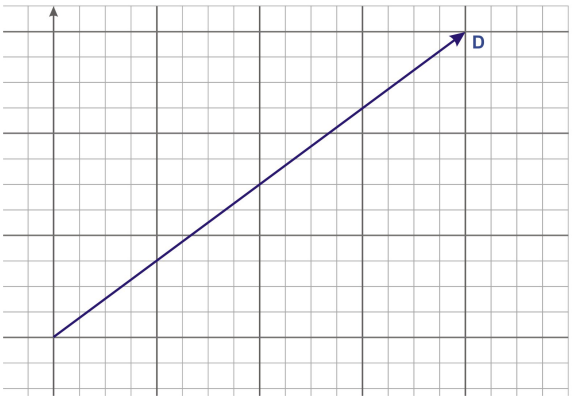
$\hat{r} = \frac{\vec{r}}{r}$

A unit vector is indicated by placing a carrot above the symbol rather than the usual vector arrow. Standard unit vectors are used to represent each of the rectangular coordinate directions: $\hat{x}, \hat{y}$, and $\hat{z}$ or $\hat{i}, \hat{j}$, and $\hat{k}$ respectively.



Using unit vector notation, the vector shown above can be written as $\vec{A} = (10 \cos 30°)\hat{x} + (10 \sin 30°)\hat{y}$ or $\vec{A} = (10 \cos 30°)\hat{i} + (10 \sin 30°)\hat{j}$

Example: What are the components of the unit vector that represents the direction of the vector shown in the diagram below?



Solution: The unit vector is defined by $\hat{r} = \frac{\vec{r}}{r}$, therefore, find the magnitude of this vector first.

$\vec{D} = \langle 3, -1.25 \rangle$

$D = \sqrt{(D_x)^2 + (D_y)^2}$

$D = \sqrt{4^2 + 3^2} = \sqrt{16 + 9} = \sqrt{25}$

$D = 5$

Then,

$\vec{D} = \frac{\vec{D}}{|D|}$

$\vec{D} = \frac{4\hat{x}+3\hat{y}}{5} = \frac{4}{5}\hat{x} + \frac{3}{5}\hat{y}$

$\vec{D} = \langle 0.8, 0.6 \rangle$

Lesson Summary

Vectors are very powerful tools used by computer programmers, mathematicians, and scientists to describe quantities which have both a size and a direction. Vectors are represented visually using arrows whose length represents the magnitude of the vector and whose orientation represents the direction of the vector. We also have two ways to represent a vector symbolically. The first identifies the vector using its magnitude and direction, $\vec{V} = 2.5 \text{ @ } 60°$. The second method uses the vector's components, which are the extension of the vector along each of the coordinate directions, $\vec{P} = 2\hat{x} + 3\hat{y}$. The arrows over the V and P identify these quantities as vectors and the carrot over the x and y identify these as unit vectors. A unit vector is a vector which has unit magnitude. We can also use a shorthand version of the component notation by writing the components in order within a set of brackets, $\vec{P} = \langle 2, 3 \rangle$.

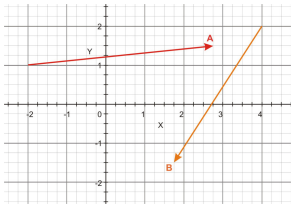
Vectors can be combined mathematically using addition and subtraction, but adding vectors is different from adding scalars. When adding scalar numbers, you just add the numbers together, for example 2 + 3 = 5. To add vector quantities, add their x-components to one another, the y-components to one another, and so on. For example, if $\vec{P} = \langle 2, 3 \rangle$ and $\vec{R} = \langle 5, 6 \rangle$, $\vec{P} + \vec{R} = \langle 7, 9 \rangle$.

Points to Consider

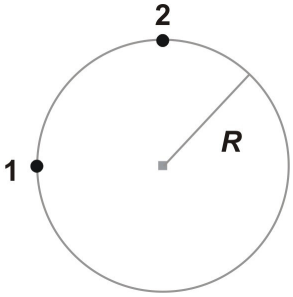
1. Can we use vectors to describe systems with more dimensions that just two?
2. Will the Pythagorean Theorem work to determine the length of a three-dimensional vector?

3. We saw in the Blue Lake problem that vectors can be used to identify paths between two points in 2D space? What are the general rules for identifying points and the distances between them?

Practice Problems



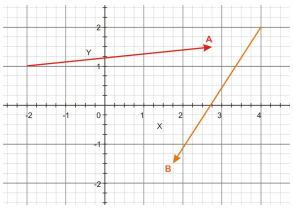
1. 1. Two vectors, $\vec{A}$ and $\vec{B}$, are shown in the diagram. Identify the $\hat{x}$ and $\hat{y}$ components of both vectors.
2. For the two vectors $\vec{A}$ and $\vec{B}$ shown in the diagram, determine the components, magnitude, and orientation angle for vector $\vec{C} = \vec{A} + \vec{B}$.
3. $\vec{A} = \langle 12, 7, 9.5 \rangle$ and $\vec{B} = \langle 8, 8, 11 \rangle$ Determine the vector $\vec{C}$ that makes the following equation true: $\vec{A} + \frac{1}{2}\vec{C} = \frac{2}{3}\vec{B}$.
4. A vector can be written as $\vec{R} = 2.74m$ @ 60°. Identify this same vector using component notation.
5. Two vectors are identified as $\vec{R} = 4.5m$ @ 20° and $\vec{R} = 6.3m$ @ 135°. Determine the magnitude and orientation angle for the sum of these two vectors.



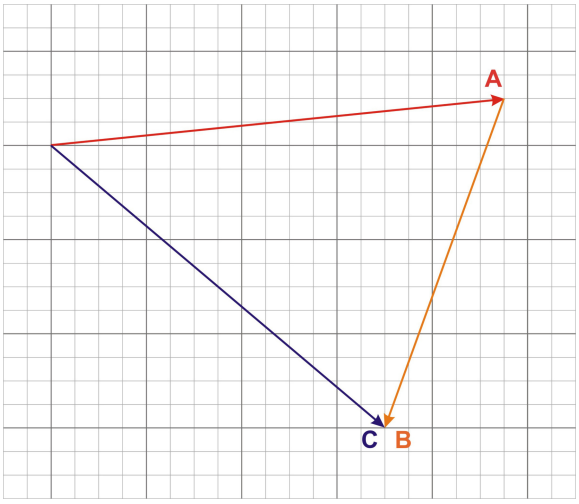
6. My neighbors recently adopted a retired racing greyhound named Flash. When Josef and Zara return home from work, they let Flash out into the back yard where she runs counter-clockwise around a circular path, similar to the one shown in the diagram. Joe recently used this circular motion to help his son Ahmed learn the difference between the vector quantity displacement and the scalar quantity distance. An object's displacement is defined as the vector beginning at the object's initial position and ending at the object's final position. The exact path does not matter to the displacement. In contrast, the distance traveled by an object does depend on the path taken in getting from its initial position to its final position. If Flash's circular path has a radius of 3.7 m, identify the displacement vector, $\vec{\Delta r}$, and the distance traveled, d, as Flash runs from point 1 to point 2.
7. Two displacement vectors have magnitudes of 2.5 cm and 6.0 cm. Is it possible to add these two vectors to obtain a vector having a magnitude of 7.5 cm? Explain why or why not using a diagram to illustrate your reasoning.
8. Three vectors are identified by the following component equations: $\vec{P} = \langle 25, 17, 32 \rangle$, $\vec{L} = \langle 14, 23, 57 \rangle$, and $\vec{Q} = \langle 49, 11, 27 \rangle$. Determine the components and magnitude of the vector $\vec{R} = 3\vec{P} + 2\vec{L} - \vec{Q}$.
9. Sasha's three puppies frequently fight over their favorite toy, a bright green Frisbee. Smoky pulls on the Frisbee with a force $\overrightarrow{F_{SonF}} = (-5.5N)\hat{x}$, Ginger pulls with a force $\overrightarrow{F_{GonF}} = (-5.2N)\hat{y}$, and Bruno pulls with a force $\overrightarrow{F_{BonF}} = (+6.0N)\hat{x}$. What is the total force acting on the Frisbee?
10. Aunt Francis frequently drives around Blue Lake to visit her friend, Rhoda, who lives at the corner

marked C in the diagram below. Determine the components of the unit vector which identifies the direction from Aunt Francis's home to Rhoda's house?

Solutions



1. Two vectors, $\vec{A}$ and $\vec{B}$, are shown in the diagram. Identify the $\hat{x}$ and $\hat{y}$ components of both vectors. The $\hat{x}$ and $\hat{y}$ components of a vector are the extensions of the vector along the $\hat{x}$ and $\hat{y}$ directions. Here we can obtain that information off of the grid in the diagram. Vector A begins at (-2, 1) and ends at (2.75, 1.5), therefore its x-component is given by A_x = 2.75 - (-2) = 4.75 and its y-component is given by A_y = 1.5 - 1 = 0.5.
   Vector B begins at (4, 2) and ends at (1.75, -1.5), therefore its x-component is given by B_x = 1.75 - 4 = -2.25 and its y-component is given by B_y = -1.5 - 2 = -3.5
2. For the two vectors $\vec{A}$ and $\vec{B}$ shown in the diagram, determine the components, magnitude, and orientation angle for vector $\vec{C} = \vec{A} + \vec{B}$. As we saw in the previous problem, $\vec{A} = \langle 4.75, 0.5 \rangle$ and $\vec{B} = \langle -2.25, -3.5 \rangle$. To add the two vectors, we add the x-components together and we add the y-components together to give $\vec{C} = \langle (4.75 + (-2.25)), (-3.5 + 0.5) \rangle = \langle 2.5, -3 \rangle$
   We can also add these two vectors graphically by positioning $\vec{A}$ and $\vec{B}$ head to tail. Vector $\vec{A}$ is the single vector that begins where $\vec{A}$ begins and ends where $\vec{B}$ ends. As you can see from the diagram, the components of vector $\vec{C}$ are $C_x = 2.5$ and $C_y = -3$.



3. $\vec{E} = \langle 12, 7, 9.5 \rangle$ and $\vec{F} = \langle 8, 8, 11 \rangle$ Determine the vector $\vec{G}$ that makes the following equation true: $\vec{E} + \frac{1}{2}\vec{G} = \frac{2}{5}\vec{F}$. Use standard algebraic techniques to solve for $\vec{G}$:
   $\frac{1}{2}\vec{G} = \frac{2}{5}\vec{F} - \vec{E}$
   $\vec{G} = \frac{4}{5}\vec{F} - 2\vec{E}$
   Remember that multiplying a vector by a scalar means multiplying each of the vector's components by that vector. Therefore,

$\frac{4}{5}\vec{F} = \left\langle \left(\frac{4}{5} * 8\right), \left(\frac{4}{5} * 8\right), \left(\frac{4}{5} * 11\right) \right\rangle = \left\langle \left(\frac{32}{5}\right), \left(\frac{32}{5}\right), \left(\frac{44}{5}\right) \right\rangle = \langle 6.4, 6.4, 8.8 \rangle$

$2\vec{E} = \langle (2 * 12), (2 * 7), (2 * 9.5) \rangle = \langle 24, 14, 19 \rangle$

Therefore

$\vec{G} = \langle (6.4 - 24), (6.4 - 14), (8.8 - 19) \rangle = \langle -17.6, -7.6, -10.2 \rangle$

4. A vector can be written as $\vec{R} = 2.74m$ @ 60°. Identify this same vector using component notation. When resolving a two-dimensional vector into components, remember that the vector itself is always the hypotenuse of a right triangle. If we define the x-axis as pointing from the origin along $= 0^o$ and the y-axis as pointing along $= 90^o$, the x-component is given by $R_x = R \cos$ and $R_y = R \sin$. In this case,

   $R_x = R \cos = (2.74\text{m}) \cos 60^o = 1.37\text{m}$

   $R_y = R \sin = (2.74\text{m}) \sin 60^o = 2.37\text{m}$

5. Two vectors are identified as $\vec{L} = 4.5m$ @ 20° and $\vec{N} = 6.3m$ @ 155°. Determine the magnitude and orientation angle for the sum of these two vectors. First we need to find the components of these two vectors. Again, let us define the x-axis as pointing from the origin along $= 0^o$ and the y-axis as pointing along $= 90^o$.

   $L_x = \text{L} \cos = (4.5\text{m}) \cos 20^o = 4.2\text{m}$

   $L_y = \text{L} \sin = (4.5\text{m}) \sin 20^o = 1.5\text{m}$

   $N_x = \text{N} \cos = (6.3\text{m}) \cos 155^o = \text{-}5.7\text{m}$

   $N_y = \text{N} \sin = (6.3\text{m}) \sin 155^o = 2.7\text{m}$

   Now we can add the components of the two vectors to give the components of their sum $\vec{J} = \vec{L} + \vec{N} = \langle (4.2 - 5.7), (1.5 + 2.7) \rangle = \langle -1.5, 4.2 \rangle$

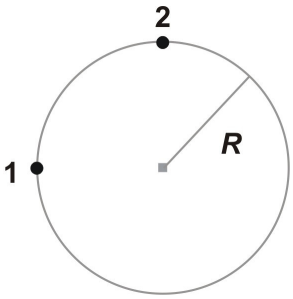
   To find the magnitude of the sum, we use the Pythagorean Theorem: $c = \sqrt{a^2 + b^2}$

   $J = \sqrt{J_x^2 + J_y^2} = \sqrt{(-1.5)^2 + (4.2)^2} = \sqrt{2.25 + 17.64} = \sqrt{19.89} = 4.46\text{m}$

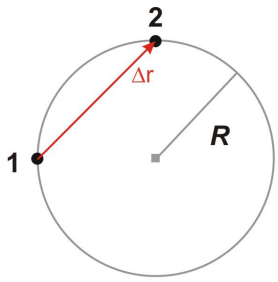
   To find the direction of the vector use right triangle trigonometry, specifically $\tan\theta = \frac{J_y}{J_x} = \frac{4.2}{-1.5} = -2.8$

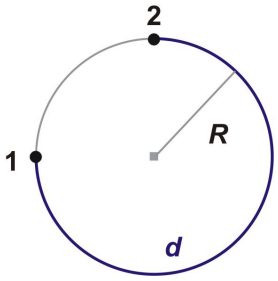
   $\theta = -109.7°$

   Note, your calculator may give you -70.3° as the angle. Many calculators only give angles from the 1st and 4th quadrant, but we know from the components that vector $\vec{J}$ is in the 2nd quadrant, but -70.3° lies within the 4th quadrant. Add 180° to the -70.3° to obtain the actual orientation angle for this vector.



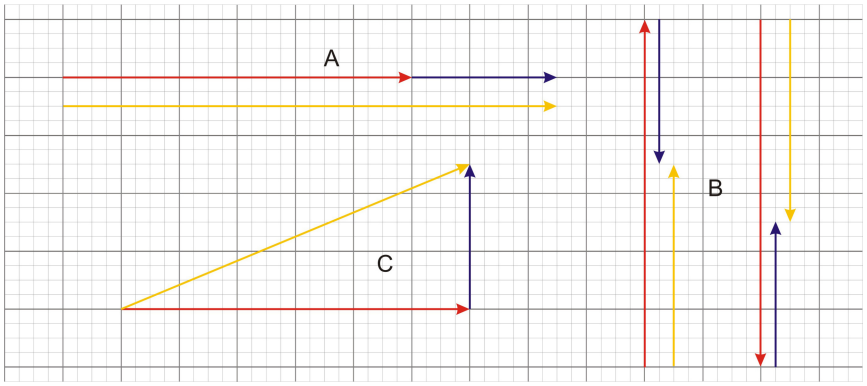
6. My neighbors recently adopted a retired racing greyhound named Flash. When Josef and Zara return home from work, they let Flash out into the back yard where she runs counter-clockwise around a circular path, similar to the one shown in the diagram. Joe recently used this circular motion to help his son Ahmed learn the difference between the vector quantity scalar quantity distance. An object's displacement is defined as the vector beginning at the object's initial position and ending at the object's final position. The exact path does not matter to the displacement. In contrast, the distance traveled by an object does depend on the path taken in getting from its initial position to its final position. If Flash's circular path has a radius of 3.7 m, identify the displacement vector, $\vec{\Delta r}$, and the distance traveled, d, as Flash runs from point 1 to point 2.





The vector beginning at point *1* and ending at point *2* is the hypotenuse of a right isosceles triangle. The length of this vector is given by the Pythagorean Theorem $\triangle r = \sqrt{R^2 + R^2} = \sqrt{2R^2} = R\sqrt{2}$
The distance traveled is the portion of the circumference traveled by Flash going counter-clockwise around the circle from point *1* to point *2*. $d = \frac{3}{4}(2\pi R) = \frac{3}{2}\pi R$

7. Two displacement vectors have magnitudes of 25 cm and 64 cm. Is it possible to add these two vectors to obtain a vector having a magnitude of 77 cm? Explain why or why not using a diagram to illustrate your reasoning. If these two numbers were scalars, then it would be impossible to add them and obtain a sum of 7.5. They are, however, vector quantities even though the problem statement did not define their directions. The largest possible sum for these two magnitudes would occur when the two vectors point in the same direction. For example if both vectors were aligned with the positive x-axis, as shown in (A) below, their sum would be 8.5 @ 0°. The smallest possible sum for these two magnitudes would occur when the two vectors point in opposite directions. For example, if one vector was parallel to the positive y-axis and the other was parallel with the negative y-axis their sum will either be +3.5 @ 90° or -3.5 @ 90°, as shown in (B) below. If the two vectors are perpendicular to one another, as shown in (C) below, the magnitude of their sum is given by the Pythagorean Theorem, $\sqrt{a^2 + b^2} = \sqrt{2.5^2 + 6^2} = \sqrt{6.25 + 36} = \sqrt{42.25} = 6.5$.



To see this in a more concrete way, take a pencil and a crayon and hold them together end to end to represent the head-to-tail addition of two vectors. The sum of the two vectors will always be the straight line that starts at the open end of the pencil and ends at the open end of the crayon.

8. Three vectors are identified by the following component equations: $\vec{P} = \langle 25, 17, 32 \rangle$, $\vec{L} = \langle 14, 23, 57 \rangle$,

and $\overrightarrow{Q} = \langle 49, 11, 27 \rangle$. Determine the components and magnitude of vector $\overrightarrow{R} = 3\overrightarrow{P} + 2\overrightarrow{L} - \overrightarrow{Q}$. The components of vector $\overrightarrow{R}$ are given by

R_x = (3*25) + (2*14) - 49 = 75 + 28 - 49 = 54

R_y = (3*17) + (2*23) - 11 = 51 + 46 - 11 = 86

R_z = (3*32) + (2*57) - 27 = 64 + 114 - 27 = 151

The magnitude of vector $\overrightarrow{R}$ is given by using the Pythagorean Theorem.

$R = \sqrt{a^2 + b^2 + c^2} = \sqrt{54^2 + 86^2 + 151^2} = \sqrt{2916 + 7396 + 22801} = \sqrt{33133} = 182$

9. Sasha's three puppies frequently fight over their favorite toy, a bright green Frisbee. Smoky pulls on the Frisbee with a force $\overrightarrow{F_{SonF}} = (-4.5N)\hat{x}$, Ginger pulls with a force $\overrightarrow{F_{GonF}} = (-5.2N)\hat{y}$, and Bruno pulls with a force $\overrightarrow{F_{BonF}} = (+6.0N)\hat{x}$. What is the total force acting on the Frisbee? Find both the magnitude and the direction.

   To find the total force on the Frisbee, add the three force vectors.

   $\overrightarrow{F_{total}} = \langle (6.0 - 4.5), (-5.2) \rangle = \langle 1.5, -5.2 \rangle N$

   $\overrightarrow{F_{total}} = \sqrt{1.5^2 + (-5.2)^2} = \sqrt{29.29} = 5.4N$

10. Aunt Francis frequently drives around Blue Lake to visit her friend, Rhoda, who lives at the corner marked C in the diagram below. Determine the components of the unit vector which identifies the direction from Aunt Francis's home to Rhoda's house?

    To find the unit vector that identifies the direction from Aunt Francis's home to Rhoda's house, we first add the three vectors $\overrightarrow{FE}$, $\overrightarrow{ED}$, and $\overrightarrow{DC}$.

    $\overrightarrow{FC} = \langle (0 + 225 + 0), (60 + 0 - 310) \rangle = \langle 225, -250 \rangle$

    Then we determine the magnitude of $\overrightarrow{FC}$.

    $FC = \sqrt{225^2 + (-250)^2} = \sqrt{113125} = 336m$

    The unit vector is defined as the ratio of the vector to its magnitude, in this case

    $\overrightarrow{FC} = \frac{\overrightarrow{FC}}{FC} = \left\langle \frac{225}{336}, \frac{250}{336} \right\rangle = \langle 0.670, 0.774 \rangle$

6.2 Vectors in Space

Learning objectives

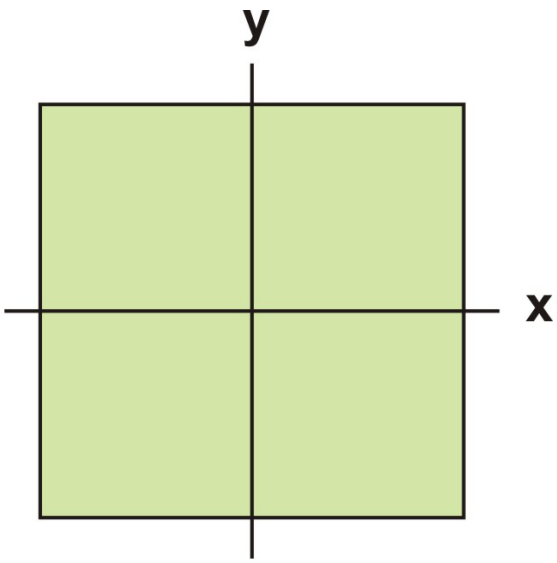
- Identify useful coordinate systems to describe a given problem.
- Identify points in 3-dimensional space using vector notation.
- Identify the displacement of an object from one point to another.
- Identify the midpoint between two points.

Introduction

In the previous section, we confined our discussion to points and vectors located in a 2-dimensional space, but these tools are not limited to describing points in a plane. In fact, mathematicians and scientists use these same techniques to describe a variety of abstract spaces as well, such as imaginary numbers, the phase relationships between a set of sinusoidal functions, or the social relationships within a high-school class. In this section, we will see that vector notation is also useful to describe the position of points or objects in a world that has at least three dimensions. Vectors can be used for more dimensions as well (but four dimensional objects are hard to draw!)

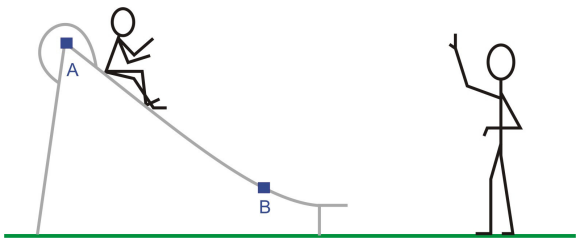
Coordinate Systems and Position Vectors in Two Dimensions

The Cartesian coordinate system used to describe a plane consists of an origin and four open axes that extend outwards from the origin towards infinity. These axes are divided into two, co-linear pairs: $+\mathbf{x}$ with $-\mathbf{x}$ and $+\mathbf{y}$ with $-\mathbf{y}$. Usually, the x-axes are oriented horizontally or left-to-right across a page and the y-axes are oriented vertically or top-to-bottom on a printed page as shown below. The green square represents the x-y plane; remember, however, that the x-y plane is not confined to this square but rather extends outward toward infinity. On most North American maps, $+\mathbf{y}$ corresponds to North and $+\mathbf{x}$ corresponds to East. We used this coordinate system to describe Allen's trip to his mother's home across Blue Lake in one of the examples of the previous sections introducing vector addition.



The choice of coordinate axes is arbitrary. Nothing about the motion of an object is limited by our choice of how to describe it. Therefore, we can choose a coordinate system that is most convenient. If we go back to the example of Blue Lake, we can identify the location of the Lake Park Lodge (L) with respect to Allen's home (P), to Aunt Frances's home (F), or to the Blue Lake Post Office (P). The point we choose as the reference position becomes our origin of coordinates (O). The position vector which identifies the Lodge's position within the reference frame is represented by an arrow starting at the origin and ending at the Lodge. This is like taking the rectangular coordinate system and putting the origin at a convenient location. The diagram below shows the position vector for the Lodge in three reference frames, with origins at F, A, and P respectively.

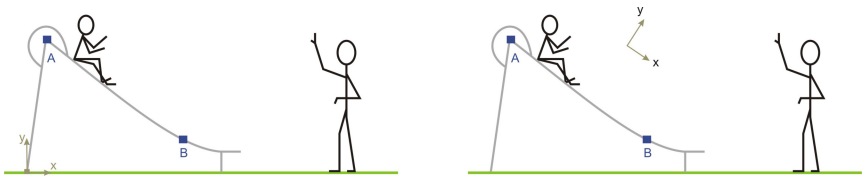
Example: The motion of an object along an inclined plane is a very common problem in introductory physics. The diagram below shows one such situation.



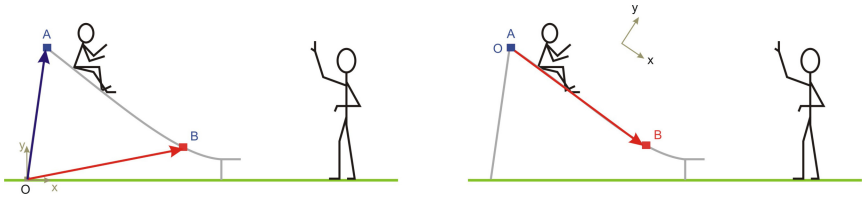
Stickman Beauford has taken his niece Brynna to the park and waves to her as she plays on the slide. Choose two coordinate systems that could be used to describe Brynna's motion and identify the position vectors for points A and B in both coordinate systems.

Solution: If we want to describe Brynna's motion as she moves from point A to point B along the slide,

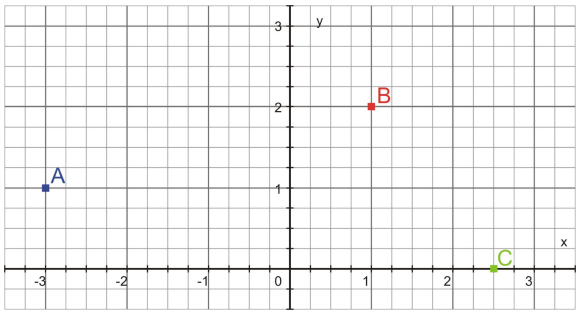
we could use a standard horizontal and vertical coordinate system with the origin at the base of the slide's ladder, but then the vector describing her motion would have components in both the x and y directions. Our mathematical description of her motion can be greatly simplified if we choose point A to be the origin and if we rotate the coordinate system such that the x-axis is parallel to the slide and the y-axis is perpendicular to the slide. Now Brynna's motion from point A to point B is only along the x-axis. Note, other choices of origin are possible.



Once we have identified an origin and coordinate axes for of our reference frames, we can use vector notation to identify the location of points A and B. The position vector for point A is the vector starting at the origin and ending at point A, $\overrightarrow{OA}$. For the standard coordinate system shown on the left above, the position vectors $\overrightarrow{OA}$ and $\overrightarrow{OB}$ are shown on the left below. For the rotated coordinate system, shown on the right above, the position vector $\overrightarrow{OA} = 0$ and $\overrightarrow{OB} = \overrightarrow{AB}$.



Example: Identify the position vectors for the three points shown on the grid below.

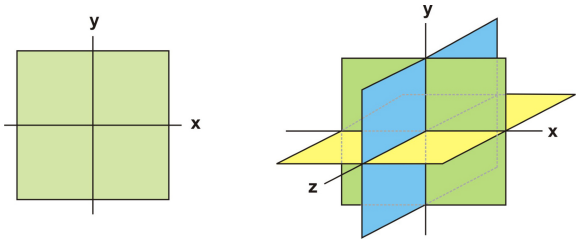


Solution: The position vectors begin at the origin, (0, 0) and end at each point:

$\overrightarrow{OA} = \langle -3, 1 \rangle$, $\overrightarrow{OB} = \langle 1, 2 \rangle$ and $\overrightarrow{OC} = \langle 2.5, 0 \rangle$

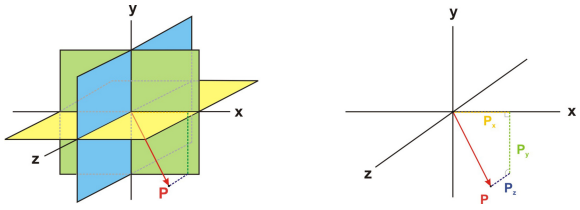
Coordinate Systems and Position Vectors in Three Dimensions

The rectangular (or Cartesian) coordinate system is used to describe a plane divided into four quadrants, as shown below left. (Note, the colored squares are used to help you visualize the space, remember that the coordinate planes actually extend outward toward infinity.)



The Cartesian coordinate system use to describe three-dimensional space consists of an origin and six open axes, +**z** and –**z** are perpendicular to the **x-y** plane. These axes define three planes which divide the space into eight quadrants as shown above right. Think of these planes as cutting space three ways: left to right, top to bottom, and front to back.

By convention, we number the four quadrants of the x-y plane in this way: points in quadrant 1 have +x and +y coordinates, those in quadrant 2 have –x and +y, those in quadrant 3 have –x and –y, and those in quadrant 4 have +x and –y. There is currently no standardized numbering system for the octants in three-dimensional space, although most people identify the region with +x, +y, and +z as the first octant. The method used to identify the octants is to indicate verbally the portion of space they occupy. For example, the first octant could also be identified as (top, front, right).



Position vectors in 3D space are still represented by arrows that begin at the origin and end at the point in question. The diagram above shows a point, P, located in the front, lower, right octant. The three components of the position vector (P_x, P_y, and P_z) are shown in the diagram. According to the Pythagorean Theorem, the magnitude of the position vector is given by:

$$|\vec{P}| = \sqrt{P_x^2 + P_y^2 + P_z^2}$$

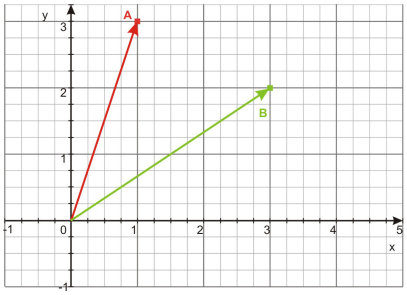
Example: Darnell was driving home from a football game in a nearby town when he swerved to avoid a deer which had run onto the road. Fortunately for Darnell, he was able to avoid hitting the deer. Unfortunately, his car ended up in the ditch beside the road. When he was unable to remove the car from the ditch by himself, he walked across a nearby field to the Tucker family farm to ask for help. The topographical map below shows Darnell's trip across the field. He traveled 300 yards south and 750 yards west from where he left his car. The map shows that he also walked uphill from an altitude of 800 feet to an altitude of 850 feet above sea level. If we treat the location of Darnell's car as the origin of coordinates, what the position vector of the Tucker farm?

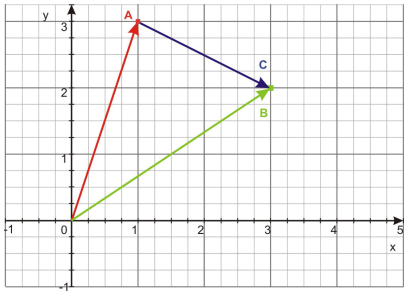
Solution: Define a coordinate system where **x** = E, **y** = N, and **z** = up. Since Darnell walked south and west from the car, the **x** and **y** coordinates of the farm are both negative. If we measure all of the distances in feet (1 yard = 3 feet), the farm's position vector can be written as $\vec{P} = \langle P_{east}, P_{north}, P_{up} \rangle = \langle -2250, -900, 50 \rangle$.

Note in this example that Darnell's walking distance was given in yards, while the elevation change was given in feet. You need to watch out for these small changes when you are solving real-life problems.

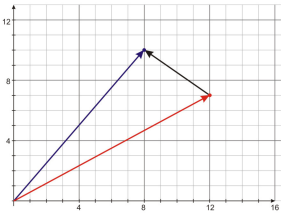
Vectors Between Two Points

The **displacement vector** represents the motion beginning at one point and ending at another. In the diagram below, vector C is the vector between points A and B. It starts at point A and ends at point B. As we saw ((in section 5.1.2)), this means that $\overrightarrow{A} + \overrightarrow{C} = \overrightarrow{B}$ and $\overrightarrow{C} = \overrightarrow{B} - \overrightarrow{A}$. In this case, $\overrightarrow{A} = \langle 1, 3\rangle$ and $\overrightarrow{B} = \langle 3, 2\rangle$, therefore $\overrightarrow{C} = \langle (3-1), (2-3)\rangle = \langle 2, -1\rangle$ matching what we can see in the diagram.





Example: Determine the coordinates and magnitude of the vector, **D**, beginning at the point $\overrightarrow{P_1} = \langle 12, 7\rangle$ and ending at $\overrightarrow{P_2} = \langle 8, 10\rangle$.



Solution: The displacement vector **D** is the difference between the two position vectors:.

$D = \left\langle P_{2x} - P_{1x}, P_{2y} - P_{1y}\right\rangle = \langle 8 - 12, 10 - 7\rangle = \langle -4, 3\rangle$.

The magnitude of the vector, **D**, can be found using the Pythagorean Theorem:

$|\overrightarrow{D}| = \sqrt{(-4)^2 + (3)^2} = \sqrt{25} = 5$.

Vector to a Point Between Two Points

Computer graphics artists frequently need to know the location of a point which lies midway between two other points. Once we know the position vectors for two discrete points, we can determine the midpoint between them using their coordinates. Specifically, the midpoint between points A and B is the "average" of the two positions, therefore the coordinates of the midpoint are given by $x_{mp} = \frac{1}{2}(x_A + x_B)$, $y_{mp} = \frac{1}{2}(y_A + y_B)$, and $z_{mp} = \frac{1}{2}(z_A + z_B)$ and the position vector for the midpoint can be written as $P_{mp} = \left\langle x_{mp}, y_{mp}, z_{mp}\right\rangle$. The vector from any other point to this midpoint can then be calculated using the method described in our discussion of displacement vectors.

Example: Determine the position vector identifying the midpoint between points $\overrightarrow{P_1} = \langle 12, 7\rangle$ and $\overrightarrow{P_2} = \langle 8, 10\rangle$.

Solution: Since these two points are located in the x-y plane, the z-coordinate of both P_1 and P_2 is 0. The x-coordinate of the midpoint is given by

$x_{mp} = \frac{1}{2}(x_A + x_B) = \frac{1}{2}(12 + 8) = 10$

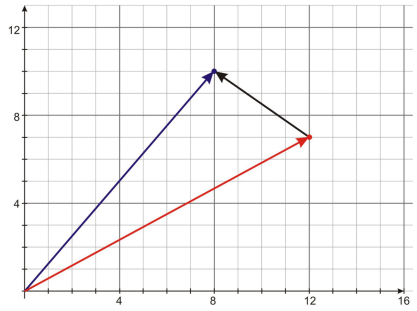
and the y-coordinate of the midpoint is given by

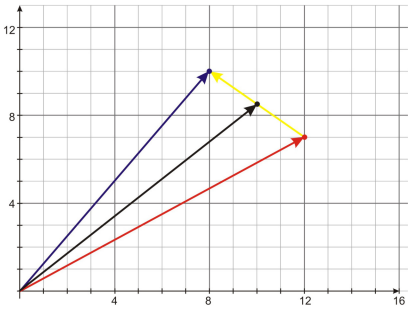
$y_{mp} = \frac{1}{2}(y_A + y_B) = \frac{1}{2}(7 + 10) = 8.5.$

Therefore the position vector for this midpoint can be written as

$P_{mp} = \langle 10, 8.5, 0 \rangle$, or if the problem context allows, you can omit the z-coordinates.

$P_{mp} = \langle 10, 8.5 \rangle$



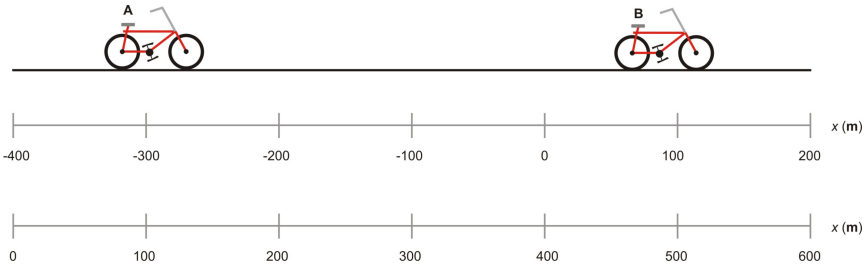


Lesson Summary

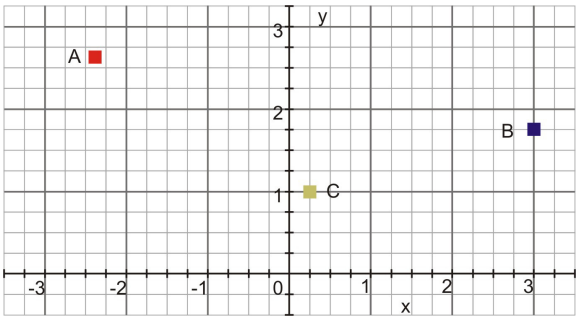
Vectors are very useful in identifying the locations of objects or points of interest. Once we define an origin of coordinates for a problem, we can identify the position of an object using the vector from that origin to the location of the object. We can also use the average of two position vectors to identify the midpoint between two points.

Practice Problems

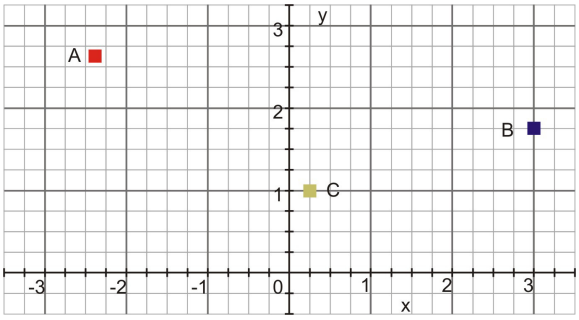
1. The diagram shows two positions of a bicycle as it moves along a long straight road. Two possible coordinate systems for the motion are shown below. Determine the position vectors in each of the two coordinate systems for the bicycle at points A and B. Then determine the displacement vector from A to B in each case. (Not drawn to scale.)



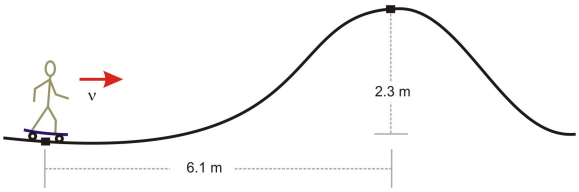
2. Identify the position vectors for the three points shown in the diagram below.



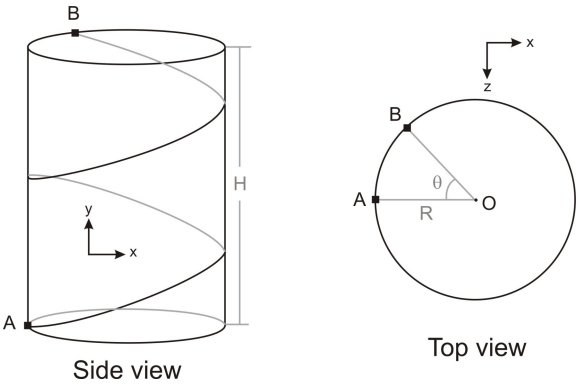
3. Identify the midpoints between points A and C and between B and C.



4. Determine the length of the displacement vectors from points A and C to the midpoint determined in the previous practice problem.
5. Zeke is enjoying an afternoon at the local skate-park. The diagram below shows his starting position and his ending position at the highest point on the new hill. Following the model of practice problem #1, choose two different coordinate systems which could describe this system. Find Zeke's initial and final position vectors in each of the two coordinate systems. Then identify the displacement vector from his starting position to his final position at the top of the hill.



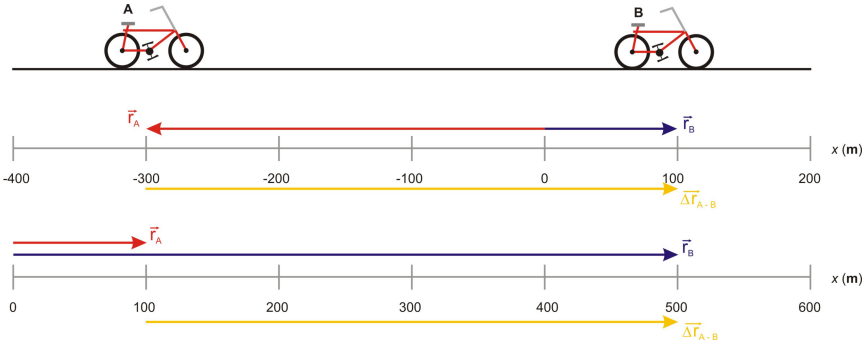
6. An architecture student designs a spiral staircase, a model of which is shown below. The staircase winds its way around a cylinder of radius 3.5 m and height 11 m. The staircase makes 1 turns progressing counter-clockwise from its beginning at point A to its end point at B. Using an origin of coordinates at the bottom center of the staircase, determine the position vectors of points A and B. Then find the displacement vector between the two points.



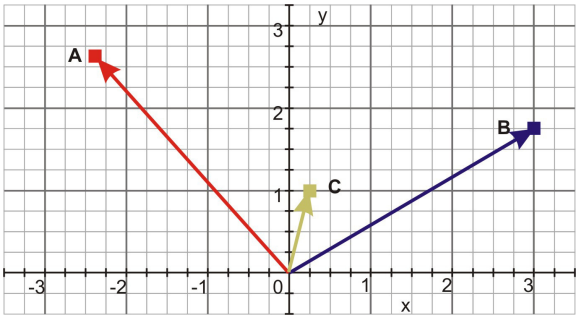
7. Identify the midpoint between points $P = (3.7, 8.4, -2.1)$ and $Q = (5.5, -1.9, -8.6)$.
8. Wilhelm and Armond head out into the marsh near their home to hunt goose for Michealmass dinner, but they disagree on the best place to set their blinds. They park their truck along the Bluffton road at an altitude of 840 m above sea level. Wilhelm heads along the riverbed, ending up 350 m west and 87 m north of the truck at an altitude of 780 m above sea level. Armond, on the other hand, heads toward another marshy area 738 m west and 92 m south of the truck at an altitude of 800 m. If the truck serves as the origin of coordinates, determine the position vector for each hunter and determine the displacement vector from Armond's position to Wilhelm's.

Solutions

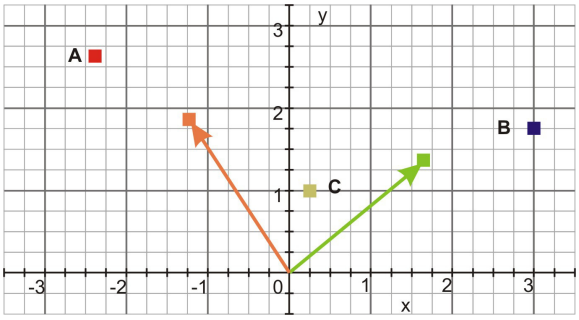
1. The diagram shows two positions of a bicycle as it moves along a long straight road. Two possible coordinate systems for the motion are shown below. Determine the position vectors in each of the two coordinate systems for the bicycle at points A and B. Then determine the displacement vector from A to B in each case. For the upper coordinate system, the position vector of the bicycle at point A is given by $\overrightarrow{r_A} = \langle -300m, 0, 0\rangle$ and that at point B is given by $\overrightarrow{r_B} = \langle 100m, 0, 0\rangle$. This gives a displacement of $\overrightarrow{\Delta r_{A-B}} = \langle (100m - (-300m)), (0-0), (0-0)\rangle = \langle 400m, 0, 0\rangle$.
   For the upper coordinate system, the position vector of the bicycle at point A is given by $\overrightarrow{r_A} = \langle 100m, 0, 0\rangle$ and that at point B is given by $\overrightarrow{r_B} = \langle 500m, 0, 0\rangle$. This gives a displacement of $\overrightarrow{\Delta r_{A-B}} = \langle (500m - 100m), (0-0), (0-0)\rangle = \langle 400m, 0, 0\rangle$.



   The position vectors for the bicycle at point A are shown in red and the position vectors for point B are shown in blue. The displacement vector between points A and B is shown in gold. As you can see, the position vectors representing this motion depend on the choice of coordinate system, but the displacement vector is independent of the coordinate system. No matter how we define the origin, the bike moves 400 m in the $+x$ direction and does not move in the y or z direction.
2. Identify the position vectors for the three points shown in the diagram below.



   $\overrightarrow{r_A} = \langle -2.63, 2.63, 0\rangle$, $\overrightarrow{r_B} = \langle 3, 1.75, 0\rangle$, $\overrightarrow{r_C} = \langle 0.25, 1, 0\rangle$
3. Identify the midpoints between points A and C and between B and C.

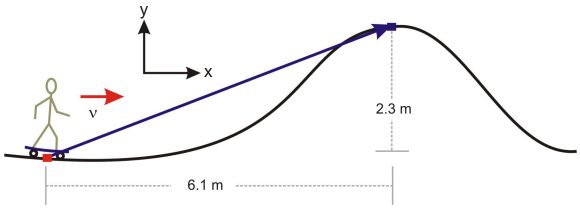


   To find the midpoint between two points, determine the average of the two positions.
   $\overrightarrow{r_{AC,mid}} = \left\langle \frac{1}{2}(-2.63 + 0.25), \frac{1}{2}(2.63 + 1), \frac{1}{2}(0 + 0)\right\rangle = \left\langle \frac{1}{2}(-2.38), \frac{1}{2}(3.63), \frac{1}{2}(0)\right\rangle = \langle -1.19, 1.815, 0\rangle$

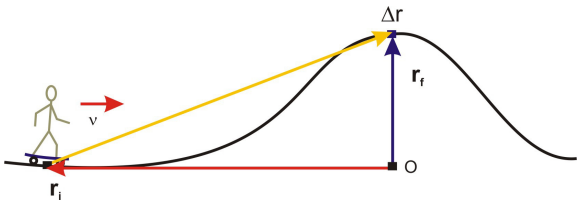
$\overrightarrow{r_{BC,mid}} = \left\langle \frac{1}{2}(3+0.25), \frac{1}{2}(1.75+1), \frac{1}{2}(0+0) \right\rangle = \left\langle \frac{1}{2}(3.25), \frac{1}{2}(2.75), \frac{1}{2}(0) \right\rangle = \langle 1.625, 1.375, 0 \rangle$

4. Determine the length of the displacement vectors from points A and C to the midpoint determined in the previous practice problem. The distance between A and the midpoint is the magnitude of the displacement vector between these two points:
$\overrightarrow{r_{AtoMid}} = \langle (-1.19-(-2.63)), (1.815-2.63), (0-0) \rangle = \langle 1.44, -0.815, 0 \rangle$
We can use the Pythagorean Theorem to calculate the length of this vector:
$|\overrightarrow{r_{AtoMid}}| = \sqrt{a^2+b^2+c^2} = \sqrt{(1.44)^2+(-0.815)^2+0^2} = \sqrt{2.0736+0.664225} = 1.655$
The distance between C and the midpoint is the magnitude of the displacement vector between these two points:
$\overrightarrow{r_{CtoMid}} = \langle (-1.19-0.25), (1.815-1), (0-0) \rangle = \langle -1.44, 0.815, 0 \rangle$
We can use the Pythagorean Theorem to calculate the length of this vector:
$|\overrightarrow{r_{CtoMid}}| = \sqrt{a^2+b^2+c^2} = \sqrt{(-1.44)^2+(0.815)^2+0^2} = \sqrt{2.0736+0.664225} = 1.655$
As you can see, the midpoint is the same distance from each of the endpoints.

5. Zeke is enjoying an afternoon at the local skate-park. The diagram below shows his starting position and his ending position at the highest point on the new hill. Following the model of practice problem #1, choose two different coordinate systems which could describe this system. Find Zeke's initial and final position vectors in each of the two coordinate systems. Then identify the displacement vector from his starting position to his final position at the top of the hill. One possible origin of coordinates is located at Zeke's starting position. In this case, the initial position vector is given by $\vec{r_i} = \langle 0m, 0m, 0m \rangle$ and his final position is given by $\vec{r_f} = \langle 6.1m, 2.3m, 0m \rangle$. Zeke's displacement is the difference between these two vectors, $\vec{\Delta r} = \vec{r_f} - \vec{r_i} = \langle 6.1m, 2.3m, 0m \rangle - \langle 0m, 0m, 0m \rangle = \langle 6.1m, 2.3m, 0m \rangle$

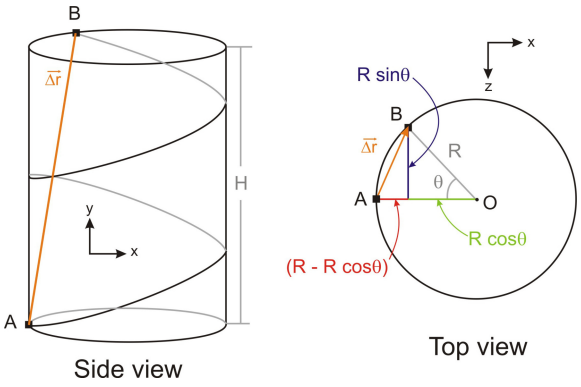


Another possible origin of coordinates is at the point marked O in the diagram below. In this case his original position is given by $\vec{r_i} = \langle -6.1m, 0m, 0m \rangle$ and his final position is given by $\vec{r_f} = \langle 0m, 2.3m, 0m \rangle$. Zeke's displacement is the difference between these two vectors, $\vec{\Delta r} = \vec{r_f} - \vec{r_i} = \langle 0m, 2.3m, 0m \rangle - \langle -6.1m, 0m, 0m \rangle = \langle 6.1m, 2.3m, 0m \rangle$



As we saw in Practice Problem #1, the position vectors representing this motion depend on the choice of coordinate system, but the displacement vector is independent of the coordinate system.

6. An architecture student designs a spiral staircase, a model of which is shown below. The staircase winds its way around a cylinder of radius 3.5 m and height 11 m. The staircase makes 1 turns progressing counter-clockwise from its beginning at point A to its end point at B. Using an origin of coordinates at the bottom center of the staircase, determine the position vectors of points A and B. Then find the displacement vector between the two points.



As you can see in the top view of the diagram, point A is directly to the left of the origin, therefore the position vector for point A is given by $\overrightarrow{r_A} = \langle -R, 0, 0\rangle = \langle -3.5m, 0, 0\rangle$.
Point B is located 11m above point A and of one turn counter-clockwise is equal to of one turn clockwise, so $= 45^o$. We can also use the geometry of the system to determine the x and z coordinates of point B: $\overrightarrow{r_B} = \langle -R\cos\theta,\ R\sin\theta,\ H\rangle = \langle(-3.5m)\cos 45°, (3.5m)\sin 45°, 11m\rangle = \langle -2.475m, 2.475m, 11m\rangle$
The displacement vector between these two points is the vector which obeys the equation: $\overrightarrow{\triangle r} = \overrightarrow{r_B} - \overrightarrow{r_A}$
$\overrightarrow{\triangle r} = \langle -2.475m, 2.475m, 11m\rangle - \langle -3.5m, 0m, 0m\rangle$
$\overrightarrow{\triangle r} = \langle(-2.475m - (-3.5m)), (2.475m - 0m), (11m - 0m)\rangle = \langle 1.025m, 2.475m, 11m\rangle$

7. Identify the midpoint between points P = (3.7, 8.4, -2.1) and Q = (5.5, -1.9, -8.6). To find the midpoint between two points, determine the average of the two positions. $M = \left(\frac{1}{2}(3.7 + 5.5), \frac{1}{2}(8.4 + (-1.9)), \frac{1}{2}(-2.1 + \left(\frac{1}{2}(9.2), \frac{1}{2}(6.5)\right), \frac{1}{2}(-10.7)\right)\right)$
$M = (4.6, 3.25, -5.35)$
8. Wilhelm and Armond head out into the marsh near their home to hunt goose for Michealmass dinner, but they disagree on the best place to set their blinds. They park their truck along the Bluffton road at an altitude of 840 m above sea level. Wilhelm heads along the riverbed, ending up 350 m west and 87 m north of the truck at an altitude of 780 m above sea level. Armond, on the other hand, heads toward another marshy area 738 m west and 92 m south of the truck at an altitude of 800 m. If the truck serves as the origin of coordinates, determine the position vector for each hunter and determine the displacement vector from Armond's position to Wilhelm's.
Define the truck as the origin of coordinates, north as the Define the truck as the origin of coordinates, north as the +y direction and east as the +x direction. Upward is the +z direction. Wilhelm's location along the river is therefore given by
$W = \langle -350m, 87m, (780m - 840m)\rangle = \langle -350m, 87m, -60m\rangle$
Armond's position at the other marsh is given by
$A = \langle -738m, -92m, (800m - 840m)\rangle = \langle -738m, -92m, -40m\rangle$
The displacement vector from Armond's position to Wilhelm's position is given by subtracting Armond's position vector from Wilhelm's position vector:
$\overrightarrow{\triangle r} = \overrightarrow{W} - \overrightarrow{A} = \langle(-350m - (-738m)), (87m - (-92m), (-60m - (-40m)))\rangle$
$\overrightarrow{\triangle r} = \langle 388m, 179m, -20m\rangle$

6.3 Dot Products

Learning objectives

- Calculate the dot product of a pair of vectors.
- Determine the angle between a pair of vectors.

- Determine the vector and scalar projections of one vector onto another.

Introduction

Previously, we discussed the multiplication of a vector by a scalar. There are actually two ways to multiply two vectors, both of which depend on the relative directions of the two vectors. One of the methods has its maximum when the two vectors are parallel; the other is maximized when the two vectors are perpendicular to one another. In this section we will look at the type of vector multiplication that gives a scalar value as the product. **The Dot Product**, also known as the scalar product, of two vectors will allow us to determine the angle between the two vectors and therefore will also help us to determine whether two vectors are parallel, perpendicular, or something in between.

Dot Product

The dot product of two vector quantities is a scalar quantity which varies as the angle between the two vectors changes. The dot product is therefore sometimes referred to as the *scalar product* of the two vectors. The maximum value for the dot product occurs when the two vectors are parallel to one another, but when the two vectors are perpendicular to one another the value of the dot product is equal to 0. Furthermore, the dot product must satisfy several important properties of multiplication. The dot product must be commutative

$\vec{A} \times \vec{B} = \vec{B} \times \vec{A}$

The dot product must also be distributive

$\vec{A} \times (\vec{B} + \vec{C}) = \vec{A} \times \vec{B} + \vec{A} \times \vec{C}$

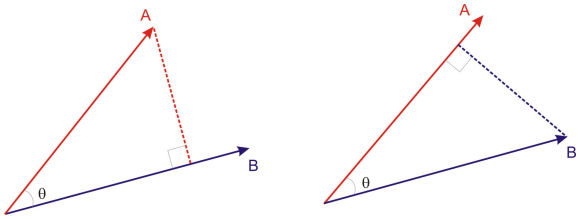
If we simply multiply the magnitudes of the two vectors together, we would fulfill the commutative property, but not the distributive property, since the magnitude

$|\vec{B} + \vec{C}| \neq |\vec{B}| + |\vec{C}|$

There are two ways to calculate the dot product. One way is to multiply the individual components. Each component of vector $\vec{A}$ is multiplied by the component of vector $\vec{B}$ which is in the same direction. Then we add the results.

$\vec{A} \times \vec{B} = A_x B_x + A_y B_y + A_z B_z + ...$

Another way to describe the process is to say that the dot product is the multiplication of one vector by the component of a second vector which is parallel to the first vector. In the diagram below are two vectors, $\boldsymbol{A}$ and $\boldsymbol{B}$. A perpendicular line has been drawn radially outward from $\boldsymbol{B}$ towards $\boldsymbol{A}$ to create a right triangle with $\boldsymbol{A}$ as the hypotenuse.



The component of $\vec{A}$ which is parallel to $\vec{B}$ is given by $A \cos$ so the second way to compute the dot product is $\vec{A} \times \vec{B} = \vec{A}(\vec{B} \cos\theta) = |\vec{A}||\vec{B}| \cos\theta$

Likewise, the component of $\vec{B}$ which is parallel to $\vec{A}$ is given by $B \cos$, so the dot product $\vec{B} \times \vec{A} =$

$|\vec{B}|(|\vec{A}| \cos\theta) = |\vec{A}||\vec{B}| \cos\theta$

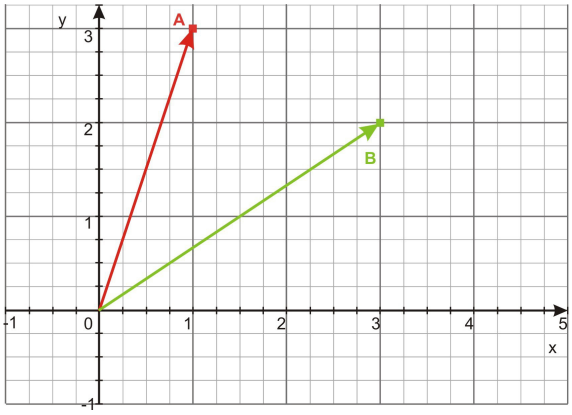
No matter which of the two vectors we project onto the other, the value of the dot product is maximized when the two vectors are parallel and zero when the two vectors are perpendicular to one another. When a vector is dotted with itself, the result is the square of the vectors magnitude since, by definition, a vector has the same direction as an equal vector.

$\vec{A} \times \vec{A} = |\vec{A}|\ (|\vec{A}| \cos\theta) = |\vec{A}||\vec{A}| \cos 0 = |\vec{A}|^2$

Also, the dot product of any vector with the zero vector is equal to zero since the magnitude of the zero vector is itself equal to 0.

Now that we have two ways to compute the dot product, we can use these two methods to solve problems about vectors. The following examples illustrate one instance of this.

Example: (a) Calculate the dot product of the two vectors shown below. (b) Then determine the angle between the two vectors.$\vec{A} = \langle 1, 3\rangle$ and $\vec{B} = \langle 3, 2\rangle$.



Solution: First we will use the components of the two vectors to determine the dot product.

$\vec{A} \times \vec{B} = A_x B_x + A_y B_y = (1 \cdot 3) + (3 \cdot 2) = 3 + 6 = 9$

Now that we know the dot product, the alternative definition of the dot product, $\vec{A} \times \vec{B} = |A||B| \cos\theta$ to find , the angle between the vectors. First find the magnitudes of the two vectors:

$|A| = \sqrt{{A_x}^2 + {A_y}^2} = \sqrt{1^2 + 3^2} = \sqrt{1+9} = \sqrt{10}$

$|B| = \sqrt{{B_x}^2 + {B_y}^2} = \sqrt{3^2 + 2^2} = \sqrt{9+4} = \sqrt{13}$

Then use these magnitudes with the cosine version of the dot product to find .

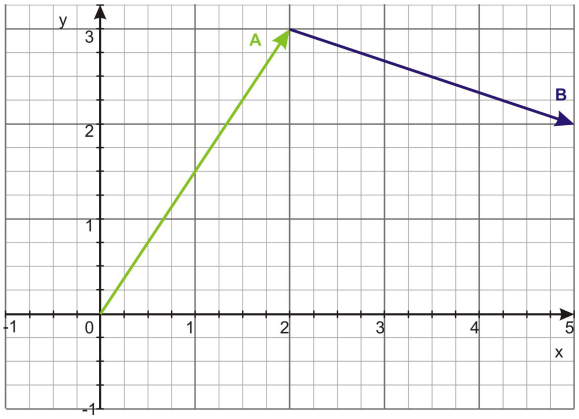
$\vec{A} \times \vec{B} = |A||B| \cos\theta$

$9 = \sqrt{10}\sqrt{13} \cos\theta$

$\cos\theta = \frac{9}{\sqrt{130}} = \frac{9}{11.4} = 0.78935$

$\theta = 37.9°$

Example: Calculate the dot product of the two vectors shown below. Then determine the angle between the two vectors.



Solution: Use the components of the two vectors to determine the dot product. Here $\overrightarrow{A} = \langle 2, 3 \rangle$ and $\overrightarrow{B} = \langle 3, -1 \rangle$.

$\overrightarrow{A} \times \overrightarrow{B} = A_x B_x + A_y B_y = (2 \cdot 3) + (3 \cdot -1) = 6 + (-3) = 3$

Now to find the angle between the vectors, first find the magnitudes of the two vectors:

$|A| = \sqrt{{A_x}^2 + {A_y}^2} = \sqrt{2^2 + 3^2} = \sqrt{4 + 9} = \sqrt{13}$

$|B| = \sqrt{{B_x}^2 + {B_y}^2} = \sqrt{3^2 + (-1)^2} = \sqrt{9 + 1} = \sqrt{10}$

Then use these magnitudes with the cosine version of the dot product to find .

$\overrightarrow{A} \times \overrightarrow{B} = |A||B| \cos \theta$

$3 = \sqrt{10}\sqrt{13} \cos \theta$

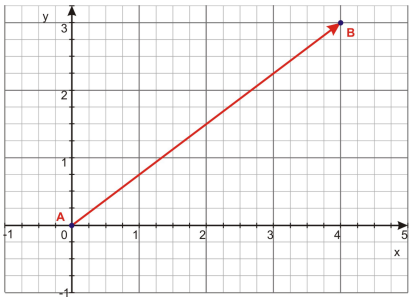
$\cos \theta = \frac{3}{\sqrt{130}} = \frac{3}{11.4} = 0.26312$

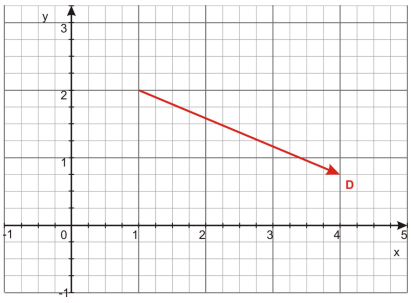
$\theta = 74.7°$

Notice that these two vectors have these same magnitudes as the vectors in Example 1, but their dot product is not the same because they are placed at different angles relative to one another.

Scalar Projections

A *scalar projection* of a vector onto a particular direction is given by the dot product of the vector with the unit vector for that direction. For example, the component notations for the vectors shown below are $AB = \langle 4, 3 \rangle$ and $D = \langle 3, -1.25 \rangle$.





The scalar projection of vector $\boldsymbol{AB}$ onto $\hat{x}$ is given by

$\overrightarrow{AB} \times \hat{x} = (4 \cdot 1) + (3 \cdot 0) + (0 \cdot 0) = 4$

The scalar projection of vector $\boldsymbol{AB}$ onto $\hat{y}$ is given by

$\overrightarrow{AB} \times \hat{y} = (4 \cdot 0) + (3 \cdot 1) + (0 \cdot 0) = 3$

And the scalar projection of vector $\boldsymbol{AB}$ onto $\hat{z}$ is given by

$\overrightarrow{AB} \times \hat{z} = (4 \cdot 0) + (3 \cdot 0) + (0 \cdot 1) = 0$

The scalar projections of $\boldsymbol{AB}$ onto the x and y directions are non-zero numbers because the vector is located in the x-y plane. The scalar projection of $\boldsymbol{AB}$ onto the z direction is equal to zero, because the z direction is perpendicular to $\boldsymbol{AB}$.

Example: The diagram below shows both vectors $\boldsymbol{AB}$ and $\boldsymbol{D}$ together on the same grid. Determine the scalar projection of vector $\boldsymbol{AB}$ onto the direction of vector $\boldsymbol{D}$.

Solution: To find the scalar projection onto the direction of another vector we need to know the unit vector in the direction of vector $\boldsymbol{D}$.

First, the components of $\overrightarrow{D}$ are

$\overrightarrow{D} = \langle 3, -1.25 \rangle$

Now the magnitude of $\overrightarrow{D}$ is

$|D| = \sqrt{(D_x)^2 + (D_y)^2} = \sqrt{3^2 + (-1.25)^2} = \sqrt{9 + 1.5625} = \sqrt{10.5625} = 3.25.$

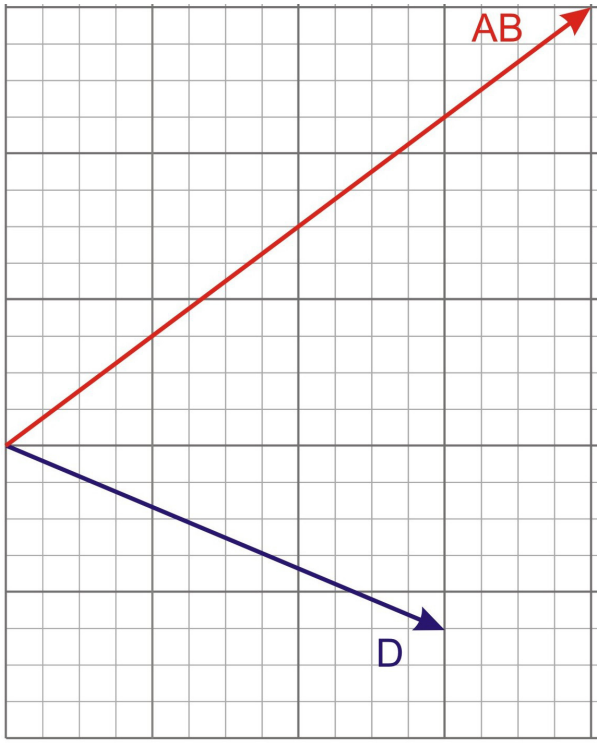
Finally, the direction vector of $\overrightarrow{D}$ is

$\overrightarrow{D} = \frac{\overrightarrow{D}}{|D|} = \frac{3\hat{x} + (-1.25)\hat{y}}{3.25} = \frac{3}{3.25}\ \hat{x} + \frac{-1.25}{3.25}\ \hat{y}.$

$\overrightarrow{D} = \langle 0.923, -0.385 \rangle$

Now we can use the dot-product to calculate the scalar projection of $\boldsymbol{AB}$ onto the direction of vector $\boldsymbol{D}$.

$\overrightarrow{AB}\ \times \overrightarrow{D} = (3 \cdot 0.923) + (4 \cdot -0.385) + (0 \cdot 0) = 2.769 + (-1.54) = 1.23$



Vector Projections

The *vector projection* of a vector onto a given direction has a magnitude equal to the scalar projection. The direction of the vector projection is the same as the unit vector of that given direction. In a previous section, we saw that when a vector $\vec{v}$ is multiplied by a scalar s, its components are given by

$\vec{sv} = \langle sv_x, sv_y, sv_z \rangle$

To calculate the vector projection of $\boldsymbol{AB}$ onto the direction of vector $\boldsymbol{D}$, use the scalar projection calculated in the previous example and the unit vector $\vec{D}$.

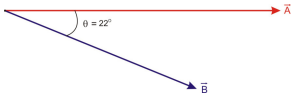
$(\vec{AB} \times \vec{D})\vec{D} = \langle (1.23 \cdot 0.923), (1.23 \cdot -0.385) \rangle = \langle 1.135, -0.474 \rangle$

Lesson Summary

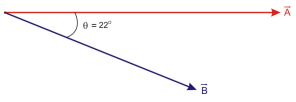
One of the two ways to multiply vector quantities is the Scalar Product. The scalar product, also known as the dot product, multiplies one vector by the component of the second vector which is parallel to the first. The result of a scalar product of two vectors is always a scalar number rather than a vector. There are two ways to calculate the dot product: $\vec{A} \times \vec{B} = |A||B| \cos\theta$ and $\vec{A} \times \vec{B} = A_xB_x + A_yB_y + A_zB_z + ...$ These two versions of the dot product can be use to determine the angle between two vectors. The dot product can also be used to determine the scalar and vector projections of one vector in the direction of the other. The scalar projection of $\vec{A}$ onto the direction of $\vec{B}$ is given by $\vec{A} \times \vec{B}$ and the vector projection of $\vec{A}$ onto the direction of $\vec{B}$ is given by $(\vec{A} \times \vec{B})\vec{B}$, where $\vec{B}$ is the unit vector in the direction of $\vec{B}$.

Practice Problems

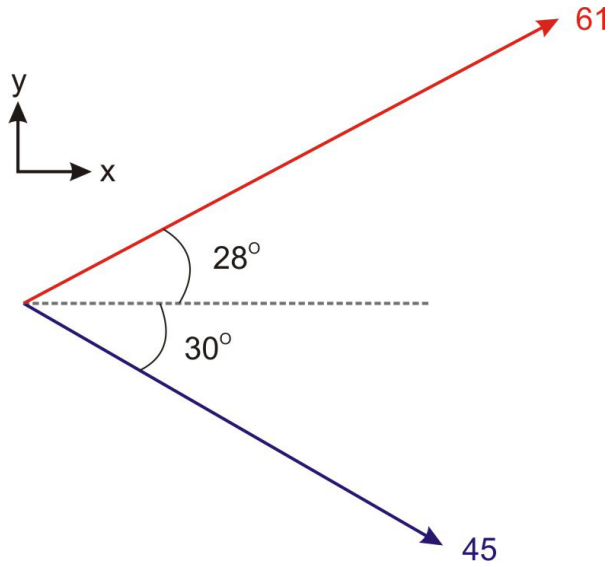
1. Determine the dot product of the two vectors $\vec{f} = \langle 3, 13, 11 \rangle$ and $\vec{g} = \langle 9, 6, 15 \rangle$.
2. Determine the dot product of the two vectors shown in the diagram below. $\vec{A} = 7cm \ @ \ 0°$ and $\vec{B} = 4cm \ @ \ -22°$.



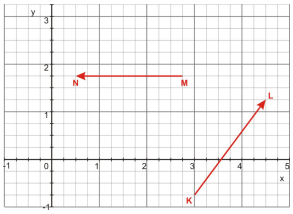
3. Determine the vector projection of $\vec{A}$ onto the direction of $\vec{B}$ and the vector projection of $\vec{B}$ onto the direction of $\vec{A}$. $\vec{A} = 7cm \ @ \ 0°$ and $\vec{B} = 4cm \ @ \ -22°$.



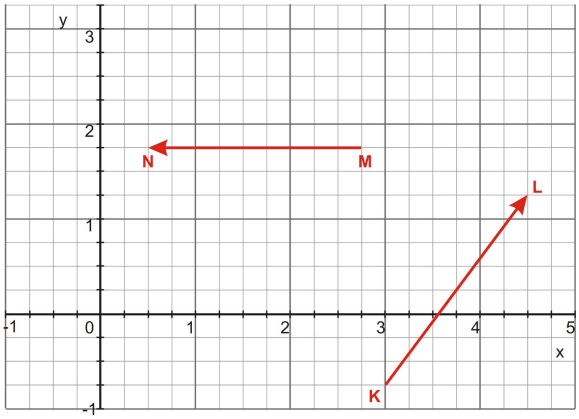
4. Determine the dot product of two vectors $\vec{E} = \langle 14, 8.5, 21 \rangle$ and $\vec{G} = \langle 15, 12.4, -3.7 \rangle$. Then determine the angle between the two vectors.
5. Determine the dot product of the following two vectors



6. Determine the scalar projection of the vector $\vec{R} = \langle 27, 39, 52 \rangle$ onto the direction of $\vec{T} = \langle 44, 26, 17 \rangle$.
7. Determine the vector projection of vector $\overrightarrow{MN}$ onto the vector $\overrightarrow{KL}$.



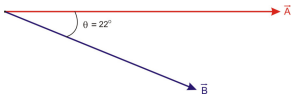
8. Determine the angle between the two vectors $\overrightarrow{MN}$ and $\overrightarrow{KL}$.



9. Determine the dot product of the two vectors $\vec{w} = \langle 85, 89, 91 \rangle$ and $\vec{h} = \langle 67, 70, 88 \rangle$, then determine the angle between the two vectors.

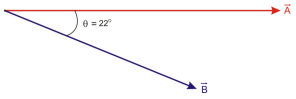
Solutions

1. Determine the dot product of the two vectors $\vec{f} = \langle 3, 13, 11 \rangle$ and $\vec{g} = \langle 9, 6, 15 \rangle$. The component form of the dot product is given by $\vec{f} \times \vec{g} = f_x g_x + f_y g_y + f_z g_z$. In this case, $\vec{f} \times \vec{g} = (3 \cdot 9) + (13 \cdot 6) + (11 \cdot 15) = 27 + 78 + 165 = 270$
2. Determine the dot product of the two vectors shown in the diagram below. $\vec{A} = 7cm \ @ \ 0°$ and $\vec{B} = 4cm \ @ \ -22°$.



The angle form of the dot product is given by $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta$. In this case, $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta = (7)(4) \cos 22 = 28 \cos 22 = 25.96\ cm^2$

3. Determine the vector projection of $\vec{A}$ onto the direction of $\vec{B}$ and the vector projection of $\vec{B}$ onto the direction of $\vec{A}$. $\vec{A} = 7cm$ @ $0°$ and $\vec{B} = 4cm$ @ $-22°$.



The vector projection of one vector onto the direction of another vector is given by $(\vec{A} \times \vec{B})\vec{B}$, where $\vec{B}$ is the unit vector in the direction of $\vec{B}$. Since it is a unit vector $\vec{B}$ has a magnitude of 1 and has the same direction as $\vec{B}$, $\vec{B} = 1$ @$-22°$. Therefore, $(\vec{A} \times \vec{B})\vec{B} = (|\vec{A}||\vec{B}| \cos\theta)\vec{B} = ((7)(1) \cos 22)$ @ $-22° = 6.49$ @ $-22°$

The vector projection of one vector onto the direction of another vector is given by $(\vec{B} \times \vec{A})\vec{A}$, where $\vec{A}$ is the unit vector in the direction of $\vec{A}$. Since it is a unit vector $\vec{A}$ has a magnitude of 1 and has the same direction as $\vec{A}$, $\vec{A} = 1$ @ $0°$. Therefore, $(\vec{B} \times \vec{A})\vec{A} = (|\vec{B}||\vec{A}| \cos\theta)\vec{A} = ((4)(1)\cos 22)$ @ $0° = 3.71$ @ $0°$

4. Determine the dot product of two vectors $\vec{E} = \langle 14, 8.5, 21\rangle$ and $\vec{G} = \langle 15, 12.4, -3.7\rangle$. Then determine the angle between the two vectors. The dot product of two vectors is defined in two ways: $\vec{A} \times \vec{B} = A_xB_x + A_yB_y + A_zB_z + ...$ and $\vec{A} \times \vec{B} = |A||B| \cos\theta$. We will use the first to calculate the dot product and then we will use that result together with the second definition to determine the angle between the two vectors.

$\vec{E} \times \vec{G} = E_xG_x + E_yG_y + E_zG_z$

$\vec{E} \times \vec{G} = (14)(15) + (8.5)(12.4) + (21)(-3.7) = 210 + 105.4 - 77.7 = 237.7$

To find the angle between the two vectors, we need to know not only the dot product of the two vectors, but also the length of each individual vector.

$|\vec{E}| = \sqrt{E_x^2 + E_y^2 + E_z^2} = \sqrt{(14)^2 + (8.5)^2 + (21)^2} = \sqrt{196 + 72.25 + 441} = \sqrt{709.25} = 26.63$

$|\vec{G}| = \sqrt{G_x^2 + G_y^2 + G_z^2} = \sqrt{(15)^2 + (12.4)^2 + (-3.7)^2} = \sqrt{225 + 153.76 + 13.69} = \sqrt{392.45} = 19.81$

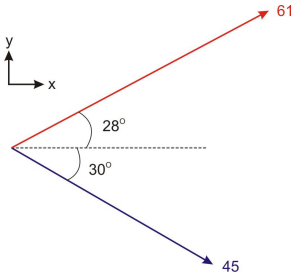
Now use the second definition of the dot product to determine the angle

$\vec{E} \times \vec{G} = |E||G|\ cos\ \theta$

$\cos\theta = \frac{\vec{E} \times \vec{G}}{|E||G|} = \frac{237.7}{(26.63)(19.81)} = \frac{237.7}{527.54} = 0.45058$

$\theta = \cos^{-1}(0.45058) = 63.2°$

5. Determine the dot product of the following two vectors.



The angle form of the dot product is given by $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta$. In this case,

$\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta = (61)(45) \cos 58 = 2745 \cos 58 = 1455$

6. Determine the scalar projection of the vector $\vec{R} = \langle 27, 39, 52\rangle$ onto the direction of $\vec{T} = \langle 44, 26, 17\rangle$.

The scalar projection of one vector onto the direction of the other is the dot product of the first vector with the unit vector representing the direction of the second vector. To calculate the scalar projection, we need to determine the unit vector in the direction of vector $\overrightarrow{T} = \langle 44, 26, 17 \rangle$. Remember that a unit vector is equal to the ratio of the vector and its magnitude, therefore we first need to calculate the length of vector $\overrightarrow{T}$.

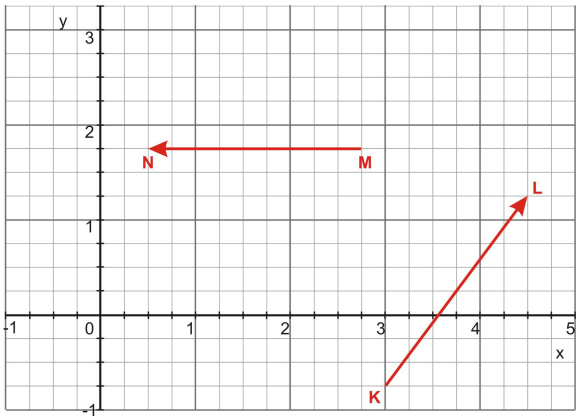
$|\overrightarrow{T}| = \sqrt{T_x^2 + T_y^2 + T_z^2} = \sqrt{(44)^2 + (26)^2 + (17)^2} = \sqrt{1936 + 676 + 289} = \sqrt{2901} = 53.86$

$\overrightarrow{T} = \frac{\overrightarrow{T}}{|\overrightarrow{T}|} = \frac{\langle 44,26,17 \rangle}{53.86} = \left\langle \frac{44}{53.86}, \frac{26}{53.86}, \frac{17}{53.86} \right\rangle = \langle 0.8169, 0.4827, 0.3156 \rangle$

Now we can calculate the scalar projection of $\overrightarrow{R}$ onto $\overrightarrow{T}$ by calculating the dot product

$\overrightarrow{R} \times \overrightarrow{T} = (27 \cdot 0.8169) + (39 \cdot 0.4827) + (52 \cdot 0.3156) = 22.0563 + 18.6253 + 16.4112 = 57.0928$

7. Determine the vector projection of vector $\overrightarrow{MN}$ onto the vector $\overrightarrow{KL}$.



The vector progression of one vector onto a second vector is the multiplication of the dot products of the two vectors and the unit vector defining the direction of the second vector. In this case, $(\overrightarrow{MN} \times \overrightarrow{KL})\overrightarrow{KL}$. First we need to identify the components of the two vectors by using the information given on the graph. In this case, $\overrightarrow{MN} = \langle -2.25, 0.0 \rangle$ and $\overrightarrow{KL} = \langle 1.5, 2, 0 \rangle$. Then we need to determine the dot product of the two vectors.

$\overrightarrow{MN} \times \overrightarrow{KL} = (MN)_x\ (MN)_x\ +\ (MN)_y\ (KL)_y\ +\ (MN)_z\ (KL)_z = (2.25)(1.5) + (2)(0) + (0)(0) = 3.375$

We also need to determine the unit vector in the direction of $\overrightarrow{KL}$.Remember that a unit vector is equal to the ratio of the vector and its magnitude, therefore we first need to calculate the length of vector $\overrightarrow{KL}$.

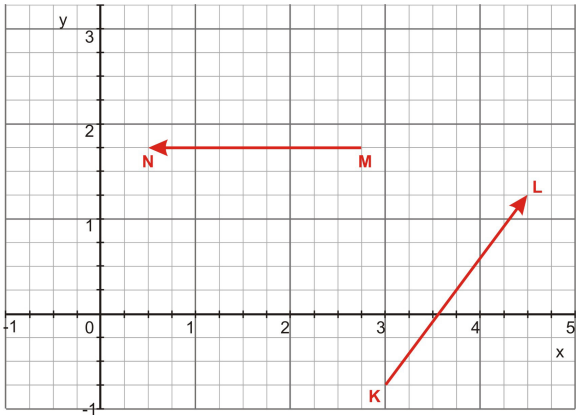
$|\overrightarrow{KL}| = \sqrt{(KL)_x^2 + (KL)_y^2 + (KL)_z^2} = \sqrt{(1.5)^2 + (2)^2 + (0)^2} = \sqrt{2.25 + 4 + 0} = \sqrt{6.25} = 2.5$

$\overrightarrow{KL} = \frac{\overrightarrow{KL}}{|\overrightarrow{KL}|} = \frac{\langle 1.5,2,0 \rangle}{2.5} = \left\langle \frac{1.5}{2.5}, \frac{2}{2.5}, \frac{0}{2.5} \right\rangle = \langle 0.6, 0.8, 0 \rangle$

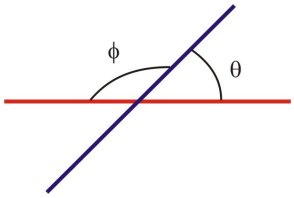
Lastly, we multiply the dot product of the two vectors by this unit vector,

$(\overrightarrow{MN} \times \overrightarrow{KL})\overrightarrow{KL} = (3.375)\langle 0.6, 0.8, 0 \rangle = \langle 2.025, 2.7, 0 \rangle$

8. Determine the angle between the vectors $\overrightarrow{MN}$ and $\overrightarrow{KL}$.



We calculated the dot product of $\overrightarrow{MN}$ and $\overrightarrow{KL}$ in the previous problem:
$\overrightarrow{MN} \times \overrightarrow{KL} = (MN)_x\ (MN)_x\ +\ (MN)_y\ (KL)_y\ +\ (MN)_z\ (KL)_z = (2.25)(1.5) + (2)(0) + (0)(0) = 3.375$
We can then use the definition $\overrightarrow{MN} \times \overrightarrow{KL} = |\overrightarrow{MN}||\overrightarrow{KL}| \cos\theta$ to determine the angle between the two vectors. But first we need to determine the magnitudes of the two vectors.
$|\overrightarrow{MN}| = \sqrt{(MN)_x^2 + (MN)_y^2 + (MN)_z^2} = \sqrt{(-2.25)^2 + (0)^2 + (0)^2} = 2.25$
$|\overrightarrow{KL}| = \sqrt{(KL)_x^2 + (KL)_y^2 + (KL)_z^2} = \sqrt{(1.5)^2 + (2)^2 + (0)^2} = \sqrt{2.25 + 4 + 0} = \sqrt{6.25} = 2.5$
$\cos\theta = \frac{\overrightarrow{MN} \times \overrightarrow{KL}}{|\overrightarrow{MN}||\overrightarrow{KL}|} = \frac{3.375}{(2.25)(2.5)} = \frac{3.375}{5.625} = 0.6$
$\theta = \cos^{-1}\ (-0.6) = 53.1°$
By looking at the diagram, we can see that the angle between these two vectors is larger than 90°. Many calculators only give the smaller of the two angles between two lines. As you can see below, both and ϕ relate the blue line to the red line.



For our problem, the calculator returned a value of 53.1°. The actual angle between the two vectors is 180° - 53.1° = 126.9° when we take into account the directions of $\overrightarrow{MN}$ and $\overrightarrow{KL}$.

9. Determine the dot product of the two vectors $\vec{w} = \langle 85, 89, 91\rangle$ and $\vec{h} = \langle 67, 70, 88\rangle$, then determine the angle between the two vectors. The component form of the dot product is given by $\vec{w} \times \vec{h} = w_x h_x + w_y h_y + w_z h_z$.
$\vec{w} \times \vec{h} = (85 \cdot 67) + (89 \cdot 70) + (91 \cdot 88) = 5695 + 6320 + 8008 = 20023$
Now we can find the angle between the two vectors using the other form of the dot-product equation: $\vec{A} \times \vec{B} = |A||B| \cos\theta$, but first we need to determine the magnitudes of the two vectors using the Pythagorean Theorem.
$|\vec{w}| = \sqrt{w_x^2 + w_y^2 + w_z^2} = \sqrt{85^2 + 89^2 + 91^2} = \sqrt{7225 + 7921 + 8281} = \sqrt{23427} = 153.1$
$|\vec{h}| = \sqrt{h_x^2 + h_y^2 + h_z^2} = \sqrt{67^2 + 70^2 + 88^2} = \sqrt{4489 + 4900 + 7744} = \sqrt{17133} = 130.9$
$\cos\theta = \frac{\vec{A} \times \vec{B}}{|A||B|} = \frac{20023}{(153.1)(130.9)} = 0.99911$
$\theta = \cos^{-1}\ 0.99911 = 2.42°$

6.4 Cross Products

Learning objectives

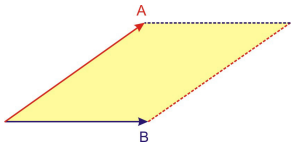
- Calculate the cross product of a pair of vectors.
- Apply the Right Hand Rule to correctly identify the direction of the vector product.
- Use the cross product to identify the direction perpendicular to a plane.

Introduction

In the last section we discussed the dot product of two vectors. In this section we will discuss the other form of vector multiplication. The dot product of two vectors produced a scalar value; the **cross product** of the same two vectors will produce a vector quantity having a direction perpendicular to the original two vectors.

Cross Product

The cross product of two vector quantities is another vector whose magnitude varies as the angle between the two original vectors changes. The cross product therefore sometimes referred to as the *vector product* of two vectors. The magnitude of the cross product represents the area of the parallelogram whose sides are defined by the two vectors, as shown in the figure below. Therefore, the maximum value for the cross product occurs when the two vectors are perpendicular to one another, but when the two vectors are parallel to one another the magnitude of the cross product is equal to zero.



Like the dot product, the cross product must satisfy several important properties of multiplication. The cross product is distributive,

$\overrightarrow{A} \times (\overrightarrow{B} + \overrightarrow{C}) = \overrightarrow{A} \times \overrightarrow{B} + \overrightarrow{A} \times \overrightarrow{C}$

But, unlike the dot product, the cross product is **anti-commutative**

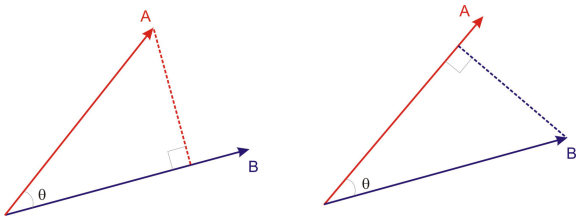
$\overrightarrow{A} \times \overrightarrow{B} = -\overrightarrow{B} \times \overrightarrow{A}$

Computing the Cross Product

The algebraic form of the cross product equation is more complicated than that for the dot product. For two 3D vectors $\overrightarrow{A}$ and $\overrightarrow{B}$,

$\overrightarrow{A} \times \overrightarrow{B} = \langle (A_2B_3 - A_3B_2), (A_3B_1 - A_1B_3), (A_1B_2 - A_2B_1) \rangle$

Another way to describe the process is to say that the cross product is the multiplication of one vector by the component of the other vector which is perpendicular to the first vector. In the diagram below are two vectors, $\boldsymbol{A}$ and $\boldsymbol{B}$. A perpendicular line has been drawn radially outward from $\boldsymbol{B}$ towards $\boldsymbol{A}$ to create a right triangle with $\boldsymbol{A}$ as the hypotenuse.



The component of $\vec{A}$ which is perpendicular to $\vec{B}$ is given by A sin so the magnitude of the cross product can be written as $|\vec{A} \times \vec{B}| = \vec{A}(\vec{B}\sin\theta) = |\vec{A}||\vec{B}| \sin\theta$

Likewise, the component of $\vec{B}$ which is parallel to $\vec{A}$ is given by B sin , so the cross product $|\vec{B} \times \vec{A}| = |\vec{B}|(|\vec{A}| \sin\theta) = |\vec{A}||\vec{B}| \sin\theta$

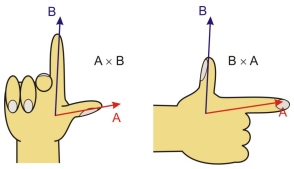
The value of the cross product is maximized when the two vectors are perpendicular and zero when the two vectors are parallel to one another. When a vector is crossed with itself, the result is the zero vector since a vector has no component perpendicular to itself.

$|\vec{A} \times \vec{A}| = |\vec{A}|(|\vec{A}| \sin\theta) = |\vec{A}||\vec{A}| \sin 0 = 0$

Also, the cross product of any vector with the zero vector is equal to zero since the magnitude of the zero vector is itself equal to 0.

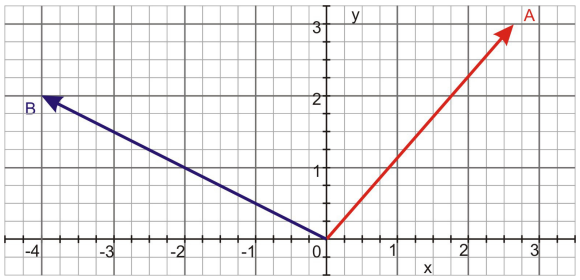
We mentioned earlier, that the direction of the cross product is perpendicular to the plane defined by the two crossed vectors. For example, the cross product of two vectors in the x-y plane will be parallel to the z-axis. This still leaves two possible directions for the cross product, though: either $+\hat{z}$ or $-\hat{z}$.

We use a **right-hand-rule** to indicate the direction of the cross product. Position the thumb and index finger of your right hand with the first vector along your thumb and the second vector along your index finger. Your middle finger, when extended perpendicular to your palm, will indicate the direction of the cross product of the two vectors.



As you can see in the diagram above, $\vec{A} \times \vec{B}$ is along $+\hat{z}$ (coming up out of the page) while $\vec{B} \times \vec{A}$ is along $-\hat{z}$ (going down into the page) and $\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$

Example: Calculate the cross product of the two vectors shown below.



Solution: Use the components of the two vectors to determine the cross product.

$\vec{A} \times \vec{B} = \langle (A_y B_z - A_z B_y), (A_z B_x - A_x B_z), (A_x B_y - A_y B_x) \rangle$

Since these two vectors are both in the x-y plane, their own z-components are both equal to 0 and the

vector product will be parallel to the z axis.

$\overrightarrow{A} \times \overrightarrow{B} = \langle[(3 \cdot 0) - (0 \cdot 2)], [(0 \cdot -4) - (2.5 \cdot 0)], [(2.5 \cdot 2) - (3 \cdot -4)]\rangle$

$\overrightarrow{A} \times \overrightarrow{B} = \langle[(0) - (0)], [(0) - (0)], [(5) - (-12)]\rangle = \langle 0, 0, (5 + 12)\rangle = \langle 0, 0, 17\rangle$

We can check our answer using the sine version of the cross product, but first we need to know the angle between the two vectors. We can use the dot product to find , following the procedure in the first Example in the previous section. First use the components to find the dot product.

$\overrightarrow{A} \times \overrightarrow{B} = A_x B_x + A_y B_y + A_z B_z = (2.5 * -4) + (3 * 2) + (0 * 0) = -10 + 6 + 0 = -4$

Then find the magnitudes of the two vectors:

$|\overrightarrow{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2} = \sqrt{2.5^2 + 3^2 + 0^2} = \sqrt{6.25 + 9 + 0} = \sqrt{15.25}$

$|\overrightarrow{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2} = \sqrt{(-4)^2 + 2^2 + 0^2} = \sqrt{16 + 4 + 0} = \sqrt{20}$

Then use these magnitudes with the cosine version of the dot product to find .

$\overrightarrow{A} \times \overrightarrow{B} = |A||B| \cos\theta$

$-4 = \sqrt{15.25}\sqrt{20} \cos\theta$

$\cos\theta = \frac{-4}{\sqrt{305}} \approx \frac{-4}{17.5} = -0.229$

$\theta = 103^\circ$

Now use the sine of this angle and the two magnitudes to determine the cross product:

$|\overrightarrow{A} \times \overrightarrow{B}| = |\overrightarrow{A}||\overrightarrow{B}| \sin\theta$

$|\overrightarrow{A} \times \overrightarrow{B}| = \sqrt{15.25}\sqrt{20} \sin 103^\circ = \sqrt{305} \sin 103^\circ = 17$

This is the same answer that we obtained from the component notation, which is good. We use the RHR to determine the direction of the vector product. If you place your thumb along vector **A** and your forefinger along vector **B**, your middle finger will point along $+\hat{z}$ and $|\overrightarrow{A} \times \overrightarrow{B}| = \langle 0, 0, 17\rangle$

The Normal Vector

We can use the cross product and the definition of the unit vector to determine the direction which is perpendicular to a plane. In the previous example, the cross product of the two vectors had a magnitude given by

$|\overrightarrow{A} \times \overrightarrow{B}| = \sqrt{15.25}\ \sqrt{20} \sin 103^\circ = \sqrt{305} \sin 103^\circ = 17$

and we used the right hand rule to show that it pointed along $+\hat{z}$ such that

$\overrightarrow{A} \times \overrightarrow{B} = \langle 0, 0, 17\rangle$

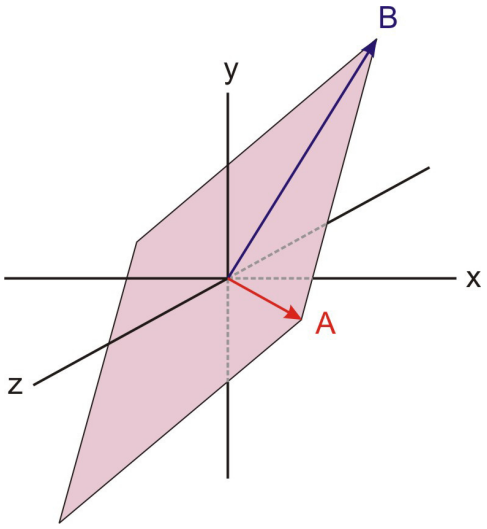
In general, we can define a normal vector, $\hat{n}$, which has a unity magnitude (i.e. magnitude equal to one) and which is perpendicular to a plane occupied by a pair of vectors, **U** and **V**.

$\hat{n} = \frac{\overrightarrow{U} \times \overrightarrow{V}}{|\overrightarrow{U} \times \overrightarrow{V}|}$

For the vectors **A** and **B** in the previous problem

$\hat{n} = \frac{\overrightarrow{A} \times \overrightarrow{B}}{|\overrightarrow{A} \times \overrightarrow{B}|} = \frac{\langle 0,0,17\rangle}{17} = \left\langle \frac{0}{17}, \frac{0}{17}, \frac{17}{17} \right\rangle = \langle 0, 0, 1\rangle$

which is in the $+\hat{z}$ direction, as shown above using the right hand rule.



Example: The diagram shows two vectors **A** and **B** which define a plane passing through the origin. Use these two vectors to determine the normal vector to this plane. $\vec{A} = \langle 3, 0, 4 \rangle$ and $\vec{B} = \langle 5, 10, 0 \rangle$

Solution: The normal vector is defined by

$\hat{n} = \frac{\vec{U} \times \vec{V}}{|\vec{U} \times \vec{V}|}$

In this case, we obtain

$\hat{n} = \frac{\vec{A} \times \vec{B}}{|\vec{A} \times \vec{B}|}$

Use the component version of the cross-product equation to find the components of $\vec{A} \times \vec{B}$

$\vec{A} \times \vec{B} = \left\langle (A_y B_z - A_z B_y), (A_z B_x - A_x B_z), (A_x B_y - A_y B_x) \right\rangle$

$\vec{A} \times \vec{B} = \langle [(0 \cdot 0) - (4 \cdot 10)], [(4 \cdot 5) - (3 \cdot 0)], [(3 \cdot 10) - (0 \cdot 5)] \rangle$

$\vec{A} \times \vec{B} = \langle (0 - 40), (20 - 0), (30 - 0) \rangle = \langle -40, 20, 30 \rangle$

Next, calculate the magnitude of the cross product, $|\vec{A} \times \vec{B}|$

$|\vec{A} \times \vec{B}| = \sqrt{(-40)^2 + 20^2 + 30^2} = \sqrt{1600 + 400 + 900} = \sqrt{2900} = 53.8516$

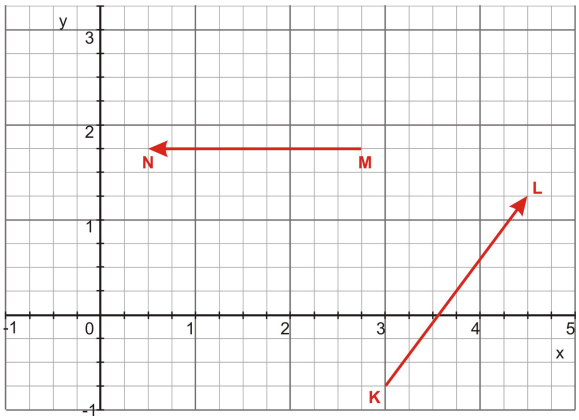
$\hat{n} = \frac{\vec{A} \times \vec{B}}{|\vec{A} \times \vec{B}|} = \frac{\langle -40, 20, 30 \rangle}{53.9} = \left\langle \frac{-40}{53.9}, \frac{20}{53.9}, \frac{30}{53.9} \right\rangle = \langle -0.743, 0.371, 0.557 \rangle$

Lesson Summary

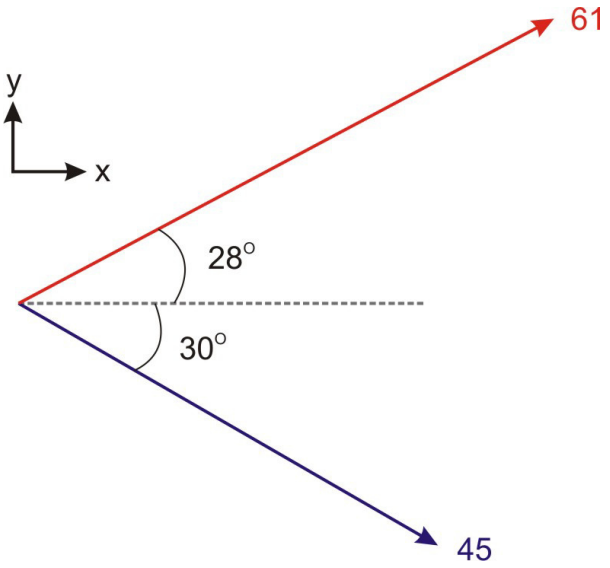
One of the two ways to multiply vector quantities is the Vector Product. The vector product, also known as the cross product, multiplies one vector by the component of the second vector which is perpendicular to the first. The result of a scalar product of two vectors is always a vector quantity which is perpendicular to the plane defined by the first two vectors. There are two ways to calculate the dot product: $\vec{A} \times \vec{B} = \left\langle (A_y B_z - A_z B_y), (A_z B_x - A_x B_z), (A_x B_y - A_y B_x) \right\rangle$ and the magnitude of the cross product is given by $|\vec{A} \times \vec{B}| = |A||B| \sin \theta$. These two versions of the cross product can be use to determine the angle between two vectors. The cross product can also be used to identify the direction perpendicular to a plane.

Practice Problems

1. Determine the magnitude and direction of the cross product $\overrightarrow{F} \times \overrightarrow{r}$ for the two vectors $\overrightarrow{F} = \langle 2, 3, 4 \rangle$ and $\overrightarrow{r} = \langle 7, 6, 5 \rangle$. Then use the cross product to determine the angle between the two vectors.
2. Determine the magnitude of the cross product of the two vectors shown below.

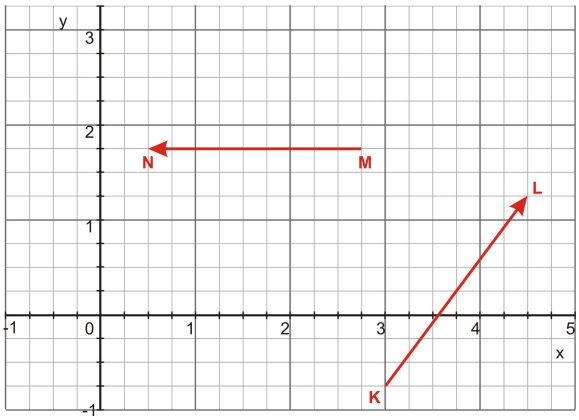


3. A plane passing through the origin is defined by the two vectors, $\overrightarrow{W} = \langle 4, 5, 2 \rangle$ and $\overrightarrow{L} = \langle 8, 1, 9 \rangle$. Determine the equation of a unit vector representing a direction perpendicular to this plane.
4. Determine the area of a parallelogram whose sides are defined by the vectors $\overrightarrow{w} = \langle 85, 89, 91 \rangle$ and $\overrightarrow{h} = \langle 67, 70, 88 \rangle$, lengths measured in centimeters.
5. Determine the magnitude of the cross product of the two vectors $\overrightarrow{f} = \langle 3, 13, 11 \rangle$ and $\overrightarrow{g} = \langle 9, 6, 15 \rangle$.
6. Determine the equation for the unit vector perpendicular to the plane defined by the two vectors $\overrightarrow{a} = \langle 2, 7, 4 \rangle$ and $\overrightarrow{b} = \langle 0, 5, 1 \rangle$.
7. Determine the area of the parallelogram whose sides are defined by $\overrightarrow{R} = \langle 27, 39, 52 \rangle$ and $\overrightarrow{T} = \langle 44, 26, 17 \rangle$, lengths measured in millimeters.
8. Determine the magnitude and direction of the cross-product of these two vectors.

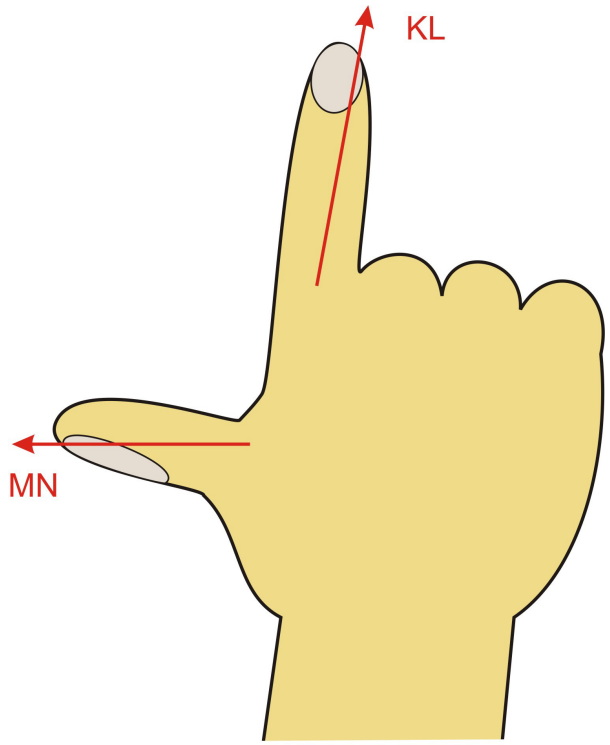


Solutions

1. Determine the cross product $\vec{F} \times \vec{r}$ for the two vectors $\vec{F} = \langle 2, 3, 4 \rangle$ and $\vec{r} = \langle 7, 6, 5 \rangle$. Then use the cross product to determine the angle between the two vectors. One of the two ways to determine the magnitude of the cross product of two vectors uses the components of the two vectors:
$\vec{F} \times \vec{r} = \left\langle (F_y r_z - F_z r_y), (F_z r_x - F_x r_z), (F_x r_y - F_y r_x) \right\rangle$
$\vec{F} \times \vec{r} = \langle (3 \cdot 5 - 4 \cdot 6), (4 \cdot 7 - 2 \cdot 5), (2 \cdot 6 - 3 \cdot 7) \rangle = \langle (15 - 24), (28 - 10), (12 - 21) \rangle$
$\vec{F} \times \vec{r} = \langle -9, 18, -9 \rangle$
Now we can use the cross product and the second definition of the cross product to determine the angle between the two vectors.
$|\vec{F} \times \vec{r}| = |F||r| \sin \theta$
We need to calculate the magnitudes of the vectors and of the cross product.
$|\vec{F}| = \sqrt{F_x^2 + F_y^2 + F_z^2} = \sqrt{2^2 + 3^2 + 4^2} = \sqrt{4 + 9 + 16} = \sqrt{29} = 5.385$
$|\vec{r}| = \sqrt{r_x^2 + r_y^2 + r_z^2} = \sqrt{7^2 + 6^2 + 5^2} = \sqrt{49 + 36 + 25} = \sqrt{110} = 10.488$
$|\vec{F} \times \vec{r}| = \sqrt{(-9)^2 + 18^2 + (-9)^2} = \sqrt{81 + 324 + 81} = \sqrt{486} = 22.0454$
$\sin \theta = \frac{|\vec{F} \times \vec{r}|}{|F||r|} = \frac{22.0454}{(5.385)(10.488)} = 0.390$
$\theta = \sin^{-1}(0.390) = 22.98°$
We can use the dot product of the two vectors to check our solution.
$\vec{F} \times \vec{r} = |\vec{F}||\vec{r}| \cos \theta$
$\vec{F} \times \vec{r} = F_x r_x + F_y r_y + F_z r_z = 2 * 7 + 3 * 6 + 4 * 5 = 14 + 18 + 20 = 52$
$\cos \theta = \frac{\vec{F} \times \vec{r}}{|\vec{F}||\vec{r}|} = \frac{52}{(5.385)(10.488)} = 0.920714$
$\theta = \cos^{-1}(0.920714) = 22.97$
This answer matches our value from the cross product to within rounding errors.
2. Determine the magnitude of the cross product of the two vectors shown below.



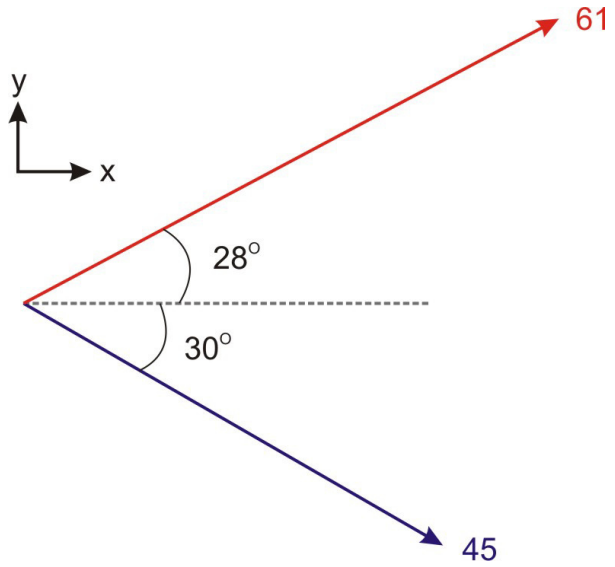
First we need to identify the components of the two vectors by using the information given on the graph. In this case, $\overrightarrow{MN} = \langle -2.25, 0.0 \rangle$ and $\overrightarrow{KL} = \langle 1.5, 2, 0 \rangle$.
$\overrightarrow{MN} \times \overrightarrow{KL} = \left\langle (MN_y KL_z - MN_z KL_y), (MN_z KL_x - MN_x KL_z), (MN_x KL_y - MN_y KL_x) \right\rangle$
$\overrightarrow{MN} \times \overrightarrow{KL} = \langle (0 \cdot 0 - 0 \cdot 2), (0 \cdot 1.5 - (-2.25) \cdot 0), ((-2.25) \cdot 2 - 0 \cdot 1.5) \rangle$
$\overrightarrow{MN} \times \overrightarrow{KL} = \langle 0 - 0, 0 - 0, -4.5 - 0 \rangle = \langle 0, 0, -4.5 \rangle$
As we can see by the components, this vector has a magnitude of 4.5 units and lies in the –z direction. We can also use the Right Hand Rule to see the direction of the cross product. As shown in the figure below, if we align the right thumb with vector MN and the right fore-finger with vector KL, the palm and extended middle-finger point in the –z direction.



3. A plane passing through the origin is defined by the two vectors, $\vec{W} = \langle 4, 5, 2\rangle$ and $\vec{L} = \langle 8, 1, 9\rangle$. Determine the equation of a unit vector representing a direction perpendicular to this plane. To solve this problem we need to use the definition of the normal vector $\hat{n} = \frac{\vec{W}\times\vec{L}}{|\vec{W}\times\vec{L}|}$, the component form of the definition of the cross product,
$\vec{W}\times\vec{L} = \left\langle (W_yL_z - W_zL_y), (W_zL_x - W_xL_z), (W_xL_y - W_yL_x)\right\rangle$. In this case, we obtain
$\vec{W}\times\vec{L} = \langle (5\cdot 9 - 2\cdot 1), (2\cdot 8 - 4\cdot 9), (4\cdot 1 - 5\cdot 8)\rangle$
$\vec{W}\times\vec{L} = \langle (45-2), (16-36), (4-40)\rangle = \langle 43, -20, -36\rangle$
We also need to know the magnitude of this cross product
$|\vec{W}\times\vec{L}| = \sqrt{x^2+y^2+z^2} = \sqrt{(43)^2+(-20)^2+(-36)^2} = \sqrt{1849+400+1296} = \sqrt{3545} = 59.54$
Now we can determine the normal vector
$\hat{n} = \frac{\vec{W}\times\vec{L}}{|\vec{W}\times\vec{L}|} = \frac{\langle 43,-20,-36\rangle}{59.54} = \left\langle \frac{43}{59.54}, \frac{-20}{59.54}, \frac{-36}{59.54}\right\rangle = \langle 0.7222, -0.3359, -0.6046\rangle$
4. Determine the area of a parallelogram whose sides are defined by the vectors $\vec{w} = \langle 85, 89, 91\rangle$ and $\vec{h} = \langle 67, 70, 88\rangle$, lengths measured in centimeters. The area of the parallelogram whose sides are defined by a pair of vectors is equal to the magnitude of the cross product of the two vectors, $|\vec{w}\times\vec{h}|$. First we need to find the cross product of the two vectors:
$\vec{w}\times\vec{h} = \left\langle (w_yh_z - w_zh_y), (w_zh_x - w_xh_z), (w_xh_y - w_yh_x)\right\rangle$
$\vec{w}\times\vec{h} = \langle (89\cdot 88 - 91\cdot 70), (91\cdot 67 - 85\cdot 88), (85\cdot 70 - 89\cdot 67)\rangle$
$\vec{w}\times\vec{h} = \langle (7832-6370), (6097-7480), (5950-5963)\rangle = \langle 1462, -1383, -13\rangle$
$|\vec{w}\times\vec{h}| = \sqrt{x^2+y^2+z^2} = \sqrt{1462^2+(-1383)^2+(13)^2} = \sqrt{4050302} \approx 2012.5$
Since the lengths of the two vectors were measured in centimeters, the area of the parallelogram is 2013 cm^2 measured to the nearest square centimeter.
5. Determine the cross product of the two vectors $\vec{f} = \langle 3, 13, 11\rangle$ and $\vec{g} = \langle 9, 6, 15\rangle$. $\vec{f}\times\vec{g} = \left\langle (f_yg_z - f_zg_y), (f_zg_x - f_xg_z), (f_xg_y - f_yg_x)\right\rangle$
$\vec{f}\times\vec{g} = \langle (13\cdot 15 - 11\cdot 6), (11\cdot 9 - 3\cdot 15), (3\cdot 6 - 13\cdot 9)\rangle$

$\vec{f} \times \vec{g} = \langle (195 - 66), (99 - 45), (18 - 117) \rangle = \langle 129, 54, -99 \rangle$

6. Determine the equation for the unit vector perpendicular to the plane defined by the two vectors $\vec{a} = \langle 2, 7, 4 \rangle$ and $\vec{b} = \langle 0, 5, 1 \rangle$. The cross product of two vectors is always perpendicular to the plane defined by the two vectors.
   $\vec{a} \times \vec{b} = \left\langle (a_y b_z - a_z b_y), (a_z b_x - a_x b_z), (a_x b_y - a_y b_x) \right\rangle$
   $\vec{a} \times \vec{b} = \langle ((7 \cdot 1) - (4 \cdot 5)), ((4 \cdot 0) - (2 \cdot 1)), ((2 \cdot 5) - (7 \cdot 0)) \rangle$
   $\vec{a} \times \vec{b} = \langle (7 - 20), (0 - 2), (10 - 0) \rangle = \langle -13, -2, 10 \rangle$
   The magnitude of this vector is given by
   $|\vec{a} \times \vec{b}| = \sqrt{x^2 + y^2 + z^2} = \sqrt{(-13)^2 + (-2)^2 + (10)^2} = \sqrt{273} = 16.5$
   Then divide the cross-product by its magnitude to obtain the unit vector.
   $\hat{n} = \frac{\vec{a} \times \vec{b}}{|\vec{a} \times \vec{b}|} = \frac{\langle -13, -2, 10 \rangle}{16.5} = \left\langle \frac{-13}{16.5}, \frac{-2}{16.5}, \frac{10}{16.5} \right\rangle$
7. Determine the area of the parallelogram whose sides are defined by $\vec{R} = \langle 27, 39, 52 \rangle$ and $\vec{T} = \langle 44, 26, 17 \rangle$, lengths measured in millimeters. The area of the parallelogram whose sides are defined by a pair of vectors is equal to the magnitude of the cross product of the two vectors, $|\vec{R} \times \vec{T}|$.
   First we need to find the cross product of the two vectors:
   $\vec{R} \times \vec{T} = \left\langle (R_y T_z - R_z T_y), (R_z T_x - R_x T_z), (R_x T_y - R_y T_x) \right\rangle$
   $\vec{R} \times \vec{T} = \langle ((39 * 17) - (52 * 26)), ((52 * 44) - (27 * 17)), ((27 * 26) - (39 * 44)) \rangle$
   $\vec{R} \times \vec{T} = \langle ((663) - (1352)), ((2288) - (459)), ((702) - (1716)) \rangle$
   $\vec{R} \times \vec{T} = \langle (663 - 1352), (2288 - 459), (702 - 1716) \rangle = \langle -689, 1829, -1014 \rangle$
   $|\vec{R} \times \vec{T}| = \sqrt{x^2 + y^2 + z^2} = \sqrt{(-689)^2 + (1829)^2 + (-1014)^2} \approx 2202$
   Since the lengths of the two vectors were measured in centimeters, the area of the parallelogram is 2202 mm^2 measured to the nearest square centimeter.
8. Determine the magnitude of the cross-product of these two vectors.



   Since we know the magnitudes of the two vectors and the angle between them, we can use the angle-version of the cross-product equation to determine the magnitude of the cross-product: $|\vec{A} \times \vec{B}| = |\vec{A}||\vec{B}| \sin\theta = (61)(45)\sin 58 = 2328$
   Since these two vectors lie in the x-y plane, the direction of the cross-product will be parallel to the z-axis.

6.5 Planes in Space

Learning objectives

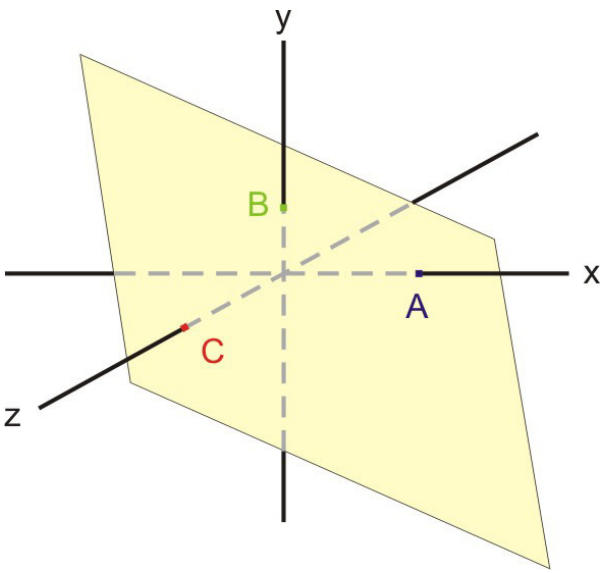
- Determine the equation of a plane given three points on the plane.
- Determine the x, y, and z intercepts of a plane.
- Determine the equation of a plane given the normal vector and a point on the plane.
- Determine the dihedral angle between two planes.

Introduction

As we have already discussed, vectors are used to indentify directions in space. We also discussed how vectors can be used to identify the orientation of a plane by identifying the direction perpendicular to that plane. In this section we will look at that calculation in reverse. Rather than determining the normal vector to a plane using two vectors which lie in that plane, we will be using the normal vector to determine the equation for the plane itself. We will also use the normal vectors to determine the intersection angle between any pair of planes.

The Equation of a Plane, Intercept Form

The diagram below shows a plane which crosses all three coordinate axes. Points A, B, and C are the locations where the plane crosses each of the coordinate axes, called *intercepts*. Their locations are given by A = (a, 0, 0), B = (0, b, 0), and C = (0, 0, c). The line segments **AB**, **BC**, and **CA** all lie in the plane. Furthermore, segment **AB** is a portion of the line of intersection between this plane and the x-y axis; segment **BC** is a portion of the line of intersection between this plane and the y-z axis; and segment **CA** is a portion of the line of intersection between this plane and the z-x axis.



The **intercept form** of the equation for a plane is given by

$1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$

This equation must hold true for all points $(x,\ y,\ z)$ on the plane. To check this equation, insert the coordinates of point A, B, or C into the equation. For points A, B, and C respectively

$1 = \frac{a}{a} + \frac{0}{b} + \frac{0}{c} = \frac{a}{a}$ $1 = \frac{0}{a} + \frac{b}{b} + \frac{0}{c} = \frac{b}{b}$ $1 = \frac{0}{a} + \frac{0}{b} + \frac{c}{c} = \frac{c}{c}$

Example: Find the intercepts of the plane given by the equation $3x + 5y - 2z - 4 = 0$.

Solution: Rewrite the equation of the plane in the format of the intercept form of the plane equation.

$3x + 5y - 2z = 4$

$1 = \frac{3}{4}x + \frac{5}{4}y + \frac{-2}{4}z$

$a = \frac{4}{3}$, $b = \frac{4}{5}$, and $c = \frac{4}{-2} = -2$. Thus the intercepts of this plane are

$\left(\frac{4}{3}, 0, 0\right)$, $\left(0, \frac{4}{5}, 0\right)$, and $(0, 0, -2)$.

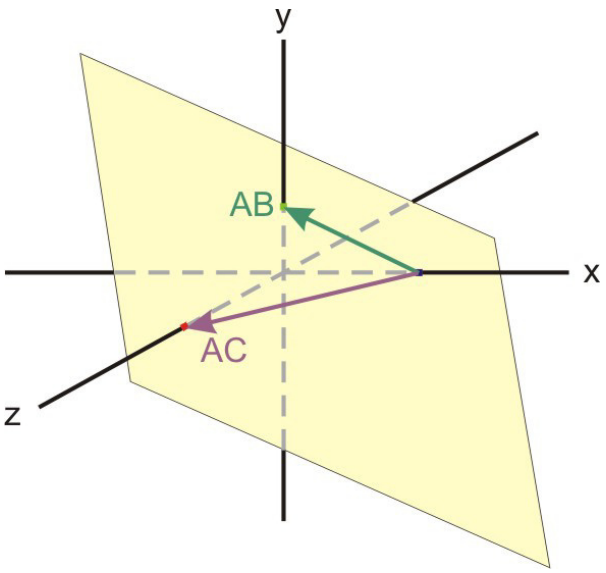
Example: A plane has intersections at (12, 0, 0), (0, 6, 0), and (0, 0, 4). Write the equation of the plane.

Solution: $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c} = \frac{x}{12} + \frac{y}{6} + \frac{z}{4}$ or $12 = x + 2y + 3z$

Identification of a Plane Using a Normal Vector

Another way to specify a plane is to know two vectors within the plane. As we showed before if two vectors lie in the same plane, the normal to that plane can be found using the cross product of the two vectors. Sometimes we are given the equations of the vectors themselves. Sometimes, however, we are only given a set of points that lie on the plane. If we know three points that lie on the plane, we can use the method developed when looking at cross products to find equations for the vectors between those points and then use those vectors to identify the plane. We only need three points to accomplish this task.

Look again at the plane we showed above. Now we will use the intercepts to determine two vectors, $\boldsymbol{AB}$ and $\boldsymbol{AC}$, which lie within the plane.



Remember that the location of points A, B, and C are given by A = (a, 0, 0), B = (0, b, 0), and C = (0, 0, c). The vectors $\boldsymbol{AB}$ and $\boldsymbol{AC}$ are then given by

$\overrightarrow{AB} = \left\langle (B_x - A_x), (B_y - A_y), (B_z - A_z) \right\rangle = \langle (0 - a), (b - 0), (0 - 0) \rangle = \langle -a, b, 0 \rangle$

$\overrightarrow{AC} = \left\langle (C_x - A_x), (C_y - A_y), (C_z - A_z) \right\rangle = \langle (0 - a), (0 - 0), (c - 0) \rangle = \langle -a, 0, c \rangle$

Now that we know two vectors in the plane, we can use the cross product of those two vectors to determine the normal to the plane.

$\overrightarrow{AB} \times \overrightarrow{AC} = \left\langle (AB_yAC_z - AB_zAC_y), (AB_zAC_x - AB_xAC_z), (AB_xAC_y - AB_yAC_x) \right\rangle$

$\overrightarrow{AB} \times \overrightarrow{AC} = \langle (bc - 0), (0 - (-ac)), (0 - (-ab)) \rangle = \langle bc, ac, ab \rangle$

Since the normal to the plane is, by definition, perpendicular to all possible vectors within a plane and since the dot product of two vectors is equal to zero for any two perpendicular vectors, we can define a plane in terms of the dot product of the normal vector with any vector, $\overrightarrow{v}$, within the plane:

$\overrightarrow{n} \times \overrightarrow{v} = 0$

Which we can also write as

$\langle n_x, n_y, n_z \rangle \times \langle (x - x_0), (y - y_0), (z - z_0) \rangle = 0$

If we compute the dot product, we obtain another equation which specifies the plane in terms of the normal vector and two points on the plane, $(x,\ y,\ z)$ and $(x_0,\ y_0,\ z_0)$.

$n_x\ (x\ -\ x_0) + n_y\ (y\ -\ y_0) + n_z\ (z\ -\ z_0) = 0$

This equation is frequently written as

$n_x x + n_y y + n_z z + d = 0$

Where

$d = -n_x x_o - n_y y_o - n_z z_o$

and the intercepts of the plane with the x, y, and z axes are given by:

$a = -\frac{d}{n_x},\ b = -\frac{d}{n_y}$, and $c = -\frac{d}{n_z}$

If you only have the normal vector and one point on the plane, first determine the vector projection of the position vector of that point onto the normal vector using the dot product. Then you will have the locations of two points on the plane and can use the normal and two points method described above.

Example: Use the equation of the plane given in the example above to determine the normal unit-vector to that plane.

Solution: The equation of the plane in the first example was $3x + 5y - 2z - 4 = 0$. Comparing this equation to $n_x x + n_y y + n_z z + d = 0$, we can see that $\overrightarrow{n} = \langle 3, 5, -2 \rangle$. To find the unit normal vector, find the magnitude of this normal vector and divide each component by the magnitude.

$|\overrightarrow{n}| = \sqrt{n_x^2 + n_y^2 + n_z^2} = \sqrt{3^2 + 5^2 + (-2)^2} = \sqrt{38}$

$\hat{n} = \left\langle \frac{3}{\sqrt{38}}, \frac{5}{\sqrt{38}}, \frac{-2}{\sqrt{38}} \right\rangle$

Example: The three points P = (3, 7, 2), Q = (1, 4, 3), and R = (2, 3, 4) define a plane. Determine the equation of the plane.

Solution: First find the vectors between two pairs of the points.

$\overrightarrow{PQ} = \left\langle (Q_x - P_x), (Q_y - P_y), (Q_z - P_z) \right\rangle = \langle (1 - 3), (4 - 7), (3 - 2) \rangle = \langle -2, -3, 1 \rangle$

$\overrightarrow{PR} = \left\langle (R_x - P_x), (R_y - P_y), (R_z - P_z) \right\rangle = \langle (2 - 3), (3 - 7), (4 - 2) \rangle = \langle -1, -4, 2 \rangle$

The cross product of these two vectors is normal to the plane.

$\overrightarrow{PQ} \times \overrightarrow{PR} = \left\langle (PQ_y PR_z - PQ_z PR_y), (PQ_z PR_x - PQ_x PR_z), (PQ_x PR_y - PQ_y PR_x) \right\rangle$

$\overrightarrow{PQ} \times \overrightarrow{PR} = \langle [(-3 \cdot 2) - (1 \cdot -4)], [(1 \cdot -1) - (-2 \cdot 2)], [(-2 \cdot -4) - (-1 \cdot -3)] \rangle$

$\overrightarrow{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = \langle [(-6) - (-4)], [(-1) - (-4)], [(8) - (3)] \rangle = \langle -2, 3, 5 \rangle$

This normal vector and one of the points will give an equation for the plane.

$n_x\ (x\ -\ P_x) + n_y\ (y\ -\ P_y) + n_z\ (z\ -\ P_z) = 0$

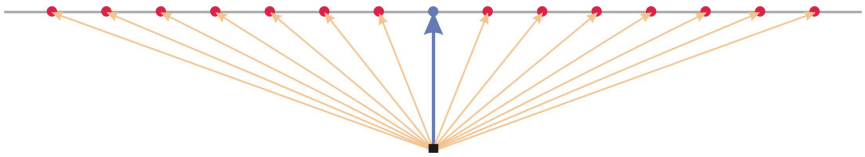
$-2(x - 3) + 3(y - 7) + 5(z - 2) = 0$

$-2x + 3y + 5z + (6 - 21 - 10) = 0$

$-2x + 3y + 5z - 25 = 0$

Determining the Distance from a Plane to the Origin

One point on the plane is unique from all the others. No matter what the orientation of the plane, this point is located closer to the origin than any other point on the plane. This means that the position vector for this point is shorter than any other point in the plane. The diagram below shows a two-dimensional projection of a plane, in grey, near a point not on the plane, in black. The position vectors to a variety of point are shown in the diagram. The position vector marked in blue is shorter than the position vectors for the other points. This shortest vector is perpendicular to the plane. You can also see that the blue line is the vector projection of any orange vector onto the perpendicular direction.



This **orthagonality** (i.e. being perpendicular) is useful for us because it means that the position vector for this special point is parallel to the normal vector. Therefore, if we know the equation for a normal vector and the position vector for any point on the plane, we can determine the location of the point on the plane closest to the origin by finding the projection of the given point's position vector onto the normal direction.

Example: The three points P = (3, 7, 2), Q = (1, 4, 3), and R = (2, 3, 4) define a plane. Determine the point on the plane which is closest to the origin.

Solution:

In the previous example, we determined the normal vector for the plane defined by points P, Q, and R: $\vec{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = \langle[(-6) - (-4)], [(-1) - (-4)], [(8) - (3)]\rangle = \langle -2, 3, 5\rangle$.

The point on the plane which is nearest to the origin can be found by determining the projection of the position vector of either of these three points onto the normal vector. Remember that the vector projection of one vector onto the direction of another, is given by the dot-product of the first vector onto the unit vector defining the direction of the second vector: $\left(\vec{A} \times \vec{B}\right)\vec{B}$.

Since we know three points on the plane, we can use one of them to solve the problem. Let's start with point P. The vector projection of $\vec{P}$ onto $\hat{n}$ is given by $\left(\vec{P} \times \hat{n}\right)\hat{n}$, so first we need to determine the unit vector $\hat{n}$ which is given by $\hat{n} = \frac{\hat{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z\rangle}{\sqrt{n_x^2+n_y^2+n_z^2}} = \frac{\langle -2,3,5\rangle}{\sqrt{(-2)^2+3^2+(5)^2}} = \frac{\langle -2,3,5\rangle}{\sqrt{38}} = \langle -0.32, 0.49, 0.81\rangle$

$\vec{P} \times \hat{n} = P_x\hat{n_x} + P_y\hat{n_y} + P_z\hat{n_z} = (3)(-0.32) + (7)(0.49) + (2)(0.81) = -0.96 + 3.43 + 1.62 = 4.09$

$\left(\vec{P} \times \hat{n}\right)\hat{n} = (4.09)\langle -0.32, 0.49, 0.81\rangle = \langle -1.3088, 2.0041, 3.3129\rangle$

Therefore, the point on the plane closest to the origin is (-1.3088, 2.0041, 3.3129).

The Dihedral Angle

The angle between two planes is called the *dihedral angle*. The angle between two planes is the same as the angle between their normal vectors. If we want to determine the dihedral angle between two planes, we identify normal vectors to the two planes, then we can use the dot-product of the two normal vectors to determine the angle between the two normals which is also the two planes. Recall $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta$

Example: The three points P = (3, 7, 2), Q = (1, 4, 3), and R = (2, 3, 4) define a plane. Determine the dihedral angle between this plane and the x-y plane.

Solution: As we saw in the example above, these three points define a plane which has a normal vector

$\vec{n} = \langle -2, 3, 5 \rangle$

The normal to the x-y plane is the unit vector $\hat{z} = \langle 0, 0, 1 \rangle$. To find the angle between these two vectors we use the fact that $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}| \cos\theta$ and that $\vec{A} \times \vec{B} = A_x B_x + A_y B_y + A_z B_z$

First find a numerical value for the dot product:

$\vec{n} \times \hat{z} = n_x z_x + n_y z_y + n_z z_z = (-2 \cdot 0) + (3 \cdot 0) + (5 \cdot 1) = 5$

$|\vec{n}| = \sqrt{n_x^2 + n_y^2 + n_z^2} = \sqrt{(-2)^2 + (3)^2 + (5)^2} = \sqrt{4 + 9 + 25} = \sqrt{38}$

$|\hat{z}| = \sqrt{z_x^2 + z_y^2 + z_z^2} = \sqrt{0^2 + 0^2 + 1^2} = 1$

Then find the cosine version of the dot product:

$\vec{n} \times \hat{z} = \sqrt{38} \cos\theta$

Now equate the two and solve for the angle,

$\vec{n} \times \hat{z} = 5 = \sqrt{38} \cos\theta$

$\theta = \cos^{-1}\left(\frac{5}{\sqrt{38}}\right) = 62.5^\circ$

Lesson Summary

Planes in space are defined by an equation which must hold true for every point on the plane. One common form of this equation is the intercept form $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$ where the intercepts are given by $\vec{x} = \langle a, 0, 0 \rangle, \vec{y} = \langle 0, b, 0 \rangle$, and $\vec{z} = \langle 0, 0, c \rangle$. Another common form of this equation can be used to determine the direction perpendicular to the plane, $n_x x + n_y y + n_z z + d = 0$ where the normal vector $\vec{n} = \langle n_x, n_y, n_z \rangle$, and the intercepts of the plane with the x, y, and z axes are given by: $\vec{n_1} \times \vec{n_2} = \cos\theta$.

$a = -\frac{d}{n_x}$, $b = -\frac{d}{n_y}$, and $c = -\frac{d}{n_z}$,

Knowing the normal vector for a plane is useful in that it allows us to determine the orientation of this plane. The dihedral angle is the angle between two planes and can be calculated using the dot product of the two normal unit-vectors. The normal vector also allows us to determine the distance from a plane to any given point in space using the scalar projection of the position vector for the given point onto the normal unit-vector defining the plane.

Practice Problems

1. Rewrite the equation of the plane $7x + 3y + z + 12 = 0$ in intercept form.
2. Determine the equation for the unit vector which is perpendicular to the plane, $7x + 3y + z + 12 = 0$.

3. Find the intercepts of the plane described by the equation $2.4x + 3.6y - 4.8z - 5.9 = 0$.
4. A plane is defined by three points having position vectors $\vec{r_1} = \langle 1, 0, -1 \rangle$, $\vec{r_2} = \langle 2, 4, 6 \rangle$, and $\vec{r_3} = \langle -3, 7, 5 \rangle$. Determine the components of the unit vector which is perpendicular to the plane passing through those points.
5. Determine the components of the unit vector which is perpendicular to the plane $12x + 23y + 14z - 5 = 0$.
6. Determine the dihedral angle between the two planes $12x + 23y + 14z - 5 = 0$ and $7x + 3y + z + 12 = 0$.
7. Determine the angle between the plane $2x - 5y + 8 - 10 = 0$ and the y-z plane.
8. The three points $\vec{P} = \langle -2, 3, 4 \rangle$, $\vec{Q} = \langle 5, -6, 7 \rangle$, and $\vec{R} = \langle 8, 9, -1 \rangle$. Determine the point on the plane which is closest to the origin.
9. Determine the point on the plane $7x + 3y + z + 12 = 0$ which is closest to the origin.

Solutions

1. Rewrite the equation of the plane $7x + 3y + z + 12 = 0$ in intercept form.The equation $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$ must be true for all points on a plane. Therefore, we should first rearrange $7x + 3y + z + 12 = 0$ into the form $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$.
   $7x + 3y + z = -12$
   $\frac{7}{-12}x + \frac{3}{-12}y + \frac{1}{-12}z = 1$
   Therefore, $a = \frac{-12}{7}$, $b = \frac{-12}{3} = -4$, and $c = \frac{-12}{1} = -12$ and the position vectors of the three intercepts are $\vec{A} = \langle -1.714, 0, 0 \rangle$, $\vec{B} = \langle 0, -4, 0 \rangle$, and $\vec{C} = \langle 0, 0, -12 \rangle$.
2. Determine the equation for the unit vector which is perpendicular to the plane, $7x + 3y + z + 12 = 0$.Comparing this equation to $n_x x + n_y y + n_z z + d = 0$, we can see that $\vec{n} = \langle 7, 3, 1 \rangle$
   $\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2 + n_y^2 + n_z^2}} = \frac{\langle 7,3,1 \rangle}{\sqrt{(7)^2 + (3)^2 + (1)^2}} = \frac{\langle 7,3,1 \rangle}{\sqrt{49+9+1}} = \frac{\langle 7,3,1 \rangle}{\sqrt{59}} = \left\langle \frac{7}{\sqrt{59}}, \frac{3}{\sqrt{59}}, \frac{1}{\sqrt{59}} \right\rangle$
3. Find the intercepts of the plane described by the equation $2.4x + 3.6y - 4.8z - 5.9 = 0$.First write the equation of the plane in intercept form, $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$.
   $2.4x + 3.6y - 4.8z = 5.9$
   $\frac{2.4}{5.9}x + \frac{3.6}{5.9}y - \frac{4.8}{5.9}z = 1$
   Therefore the x-intercept is $\left\langle \frac{5.9}{2.4}, 0, 0 \right\rangle$, the y-intercept is $\left\langle 0, \frac{5.9}{3.6}, 0 \right\rangle$, and the z-intercept is $\left\langle 0, 0, \frac{5.9}{4.8} \right\rangle$.
4. A plane is defined by three points having position *vectors* $\vec{r_1} = \langle 1, 0, -1 \rangle$, $\vec{r_2} = \langle 2, 4, 6 \rangle$, and $\vec{r_3} = \langle -3, 7, 5 \rangle$. Determine the components of the unit vector which is perpendicular to the plane passing through those points.As we saw in a previous section, the cross-product determines the direction perpendicular to a pair of vectors. Therefore we can use these three points to define two vectors in the same plane. The vector from point 1 to point 2 is given by subtracting vector 2 from vector 1:
   $\overrightarrow{r_{1-2}} = \vec{r_1} - \vec{r_2} = \langle 1, 0, -1 \rangle - \langle 2, 4, 6 \rangle = \langle 1 - 2, 0 - 4, -1 - 6 \rangle = \langle -1, -4, -7 \rangle$
   Likewise, the vector from point 1 to point 3 is given by subtracting vector 3 from vector 1:
   $\overrightarrow{r_{1-3}} = \vec{r_1} - \vec{r_3} = \langle 1, 0, -1 \rangle - \langle -3, 7, 5 \rangle = \langle 1 - (-3), 0 - 7, -1 - 5 \rangle = \langle 4, -7, -6 \rangle$
   Now we can use the cross-product of the two vectors in the plane to determine a vector which is perpendicular to that plane,
   $\vec{n} = \overrightarrow{r_{1-2}} \times \overrightarrow{r_{1-3}} = \left\langle (r_{1-2y} r_{1-3z} - r_{1-2z} r_{1-3y}), (r_{1-2z} r_{1-3x} - r_{1-2x} r_{1-3z}), (r_{1-2x} r_{1-3y} - r_{1-2y} r_{1-3x}) \right\rangle$
   $\vec{n} = \overrightarrow{r_{1-2}} \times \overrightarrow{r_{1-3}} = \langle ((-4)(-6) - (-7)(-7)), ((-7)(4) - (-1)(-6)), ((-1)(-7) - (-4)(4)) \rangle$
   $\vec{n} = \overrightarrow{r_{1-2}} \times \overrightarrow{r_{1-3}} = \langle ((24) - (49), ((-28) - (6)), ((7) - (-16)) \rangle = \langle -25, -34, 23 \rangle$
   Now use the definition of the unit vector to complete the problem.
   $\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2 + n_y^2 + n_z^2}} = \frac{\langle -25, -34, 23 \rangle}{\sqrt{(-25)^2 + (-34)^2 + (23)^2}} = \frac{\langle -25, -34, 23 \rangle}{\sqrt{625+1156+529}} = \frac{\langle -25, -34, 23 \rangle}{\sqrt{2310}}$
   $\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle -25, -34, 23 \rangle}{48.06} = \langle -0.5202, -0.7074, 0.4785 \rangle$
5. Determine the components of the unit vector which is perpendicular to the plane $12x + 23y + 14z -$

$5 = 0$.Comparing this equation to $n_x x + n_y y + n_z z + d = 0$, we can see that $\vec{n} = \langle 12, 23, 14 \rangle$.
Now we can use the definition of the unit vector to complete the problem.
$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2+n_y^2+n_z^2}} = \frac{\langle 12,23,14 \rangle}{\sqrt{12^2+23^2+14^2}} = \frac{\langle 12,23,14 \rangle}{\sqrt{869}} = \frac{\langle 12,23,14 \rangle}{29.5} = \left\langle \frac{12}{29.5}, \frac{23}{29.5}, \frac{14}{29.5} \right\rangle$

6. Determine the dihedral angle between the two planes $12x + 23y + 14z - 5 = 0$ and $7x + 3y + z + 12 = 0$.The dihedral angle is defined as the angle between two planes. This angle is also equal to the angle between the normals to the two planes. In two of the previous problems we determined the unit vectors which are perpendicular to these two planes $\vec{n_1} = \left\langle \frac{12}{29.5}, \frac{23}{29.5}, \frac{14}{29.5} \right\rangle$ and $\vec{n_2} = \left\langle \frac{7}{\sqrt{59}}, \frac{3}{\sqrt{59}}, \frac{1}{\sqrt{59}}, \right\rangle$.
We can then use the dot-product of these two normal vectors to determine the angle between the two. The dot-product is defined as $\vec{A} \times \vec{B} = A_x B_x + A_y B_y + A_z B_z + ...$ and as $\vec{A} \times \vec{B} = |A||B| \cos\theta$. First, we need to find the component version of the dot product and the magnitudes of the two normal vectors.
$\vec{n_1} \times \vec{n_2} = \vec{n_{1x}}\vec{n_{2x}} + \vec{n_{1y}}\vec{n_{2y}} + \vec{n_{1z}}\vec{n_{2z}} = \frac{12.7}{29.5\sqrt{59}} + \frac{23.3}{29.5\sqrt{59}} + \frac{14.1}{29.5\sqrt{59}}$
$\vec{n_1} \times \vec{n_2} = \frac{12 \cdot 7}{29.5\sqrt{59}} + \frac{23 \cdot 3}{29.5\sqrt{59}} + \frac{14 \cdot 1}{29.5\sqrt{59}} = \frac{119}{226.6} + \frac{69}{226.6} + \frac{14}{226.6} = \frac{202}{226.6} = 0.891$
Since these two vectors are unit-vectors, their magnitudes are both equal to 1.
$\cos\theta = \frac{\vec{n_1} \times \vec{n_2}}{|\vec{n_1}||\vec{n_2}|} = \frac{0.891}{(1)(1)} = 0.891$
$\theta = cos^{-1} 0.891 = 27.0°$

7. Determine the angle between the plane $2x - 5y + 8 - 10 = 0$ and the y-z plane.The dihedral angle is defined as the angle between the two planes and is also equal to the angle between the two normal unit vectors. In this case, we already know the normal unit vector for the y-z plane, $\hat{x} = \langle 1, 0, 0 \rangle$. We still need to determine, however, the unit vector for the plane $2x - 5y + 8z - 10 = 0$.
Comparing this equation to $n_x x + n_y y + n_z z + d = 0$, we can see that $\vec{n} = \langle 2, -5, 8 \rangle$.
Now we can use the definition of the unit vector
$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2+n_y^2+n_z^2}} = \frac{\langle 2,-5,8 \rangle}{\sqrt{2^2+(-5)^2+8^2}} = \frac{\langle 2,-5,8 \rangle}{\sqrt{4+25+64}} = \frac{\langle 2,-5,8 \rangle}{9.64} = \left\langle \frac{2}{9.64}, \frac{-5}{9.64}, \frac{8}{9.64} \right\rangle$
The angle between the two planes is equal to the angle between the two normal vectors.
We can then use the dot-product of these two normal vectors to determine the angle between the two. The dot-product is defined as $\vec{A} \times \vec{B} = A_x B_x + A_y B_y + A_z B_z + ...$ and as $\vec{A} \times \vec{B} = |A||B| \cos\theta$. First, we need to find the component version of the dot product and the magnitudes of the two normal vectors.
$\vec{n_1} \times \vec{n_2} = \vec{n_{1x}}\vec{n_{2x}} + \vec{n_{1y}}\vec{n_{2y}} + \vec{n_{1z}}\vec{n_{2z}} = \frac{2 \cdot 1}{9.64} + \frac{-5 \cdot 0}{9.64} + \frac{8 \cdot 0}{9.64} = \frac{2}{9.64} = 0.2074$
Since these two vectors are unit-vectors, their magnitudes are both equal to 1.
$\cos\theta = \frac{\vec{n_1} \times \vec{n_2}}{|\vec{n_1}||\vec{n_2}|} = \frac{0.2074}{(1)(1)} = 0.2074$
$\theta = cos^{-1} 0.2074 = 16.18°$

8. The three points $\vec{P} = \langle -2, 3, 4 \rangle$, $\vec{Q} = \langle 5, -6, 7 \rangle$, and $\vec{R} = \langle 8, 9, -1 \rangle$ identify a plane. Determine the point on the plane which is closest to the origin. The point on the plane nearest to the origin can be found by determining the projection of the position vector of one of these three points onto the normal vector. Remember that the vector projection of one vector onto the direction of another is given by the dot-product of the first vector onto the unit vector defining the direction of the second vector: $\left(\vec{P} \times \hat{n}\right)\hat{n}$.
We can use the position vectors for the three points to determine two vectors within the plane. Once we have those two vectors, their cross-product will define the direction normal to the plane. First find the two equations in the plane:
$\vec{A} = \vec{Q} - \vec{P} = \langle 5, -6, 7 \rangle - \langle -2, 3, 4 \rangle = \langle 7, -9, 3 \rangle$
$\vec{B} = \vec{R} - \vec{P} = \langle 8, 9, -1 \rangle - \langle -2, 3, 4 \rangle = \langle 10, 6, -5 \rangle$
Now determine the cross product of the two vectors
$\vec{n} = \vec{A} \times \vec{B} = \left\langle (A_y B_z - A_z B_y), (A_z B_x - A_x B_z), (A_x B_y - A_y B_x) \right\rangle$
$\vec{n} = \vec{A} \times \vec{B} = \langle (45 - 18), (30 - (-35)), (42 + 90) \rangle$

$\vec{n} = \vec{A} \times \vec{B} = \langle 27, 65, 132 \rangle$
Now we need to determine the unit vector associated with this normal vector
$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2 + n_y^2 + n_z^2}} = \frac{\langle 27,65,132 \rangle}{\sqrt{27^2 + (65)^2 + (132)^2}} = \frac{\langle 27,65,132 \rangle}{\sqrt{22378}}$
$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \langle 0.181, 0.435, 0.882 \rangle$
Now we determine the vector progression of one of the three initial position vectors onto the direction of this normal unit-vector: $\left(\vec{P} \times \hat{n}\right)\hat{n}$. Remember that the dot product is given by $\vec{A} \times \vec{B} = A_x B_x + A_y B_y + A_z B_z + ...$
$\left(\vec{P} \times \hat{n}\right)\hat{n} = (-2(0.181) + 3(0.435) + 4(0.882)) \langle 0.181, 0.435, 0.882 \rangle$
$\left(\vec{P} \times \hat{n}\right)\hat{n} = (4.471) \langle 0.181, 0.435, 0.882 \rangle$
$\left(\vec{P} \times \hat{n}\right)\hat{n} = (4.471) \langle 0.181, 0.435, 0.882 \rangle = \langle 0.809, 1.945, 3.943 \rangle$

9. Determine the point on the plane $7x + 3y + z + 12 = 0$ which is closest to the origin.The point on the plane nearest to the origin can be found by determining the projection of the position vector of any point on the plane onto the normal vector. The vector projection of one vector onto the direction of another is given by the dot-product of the first vector onto the unit vector defining the direction of the second vector: $\left(\vec{P} \times \hat{n}\right)\hat{n}$.
In this case, we can determine a normal vector using the equation of the plane. Comparing $7x + 3y + z + 12 = 0$ to the generic equation $n_x x + n_y y + n_z z + d = 0$, we can see that $\vec{n} = \langle 7, 3, 1 \rangle$ and
$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{\langle n_x, n_y, n_z \rangle}{\sqrt{n_x^2 + n_y^2 + n_z^2}} = \frac{\langle 7,3,1 \rangle}{\sqrt{(7)^2 + (3)^2 + (1)^2}} = \frac{\langle 7,3,1 \rangle}{\sqrt{49+9+1}} = \frac{\langle 7,3,1 \rangle}{\sqrt{59}} = \left\langle \frac{7}{\sqrt{59}}, \frac{3}{\sqrt{59}}, \frac{1}{\sqrt{59}} \right\rangle$
We also need to know the location of a point on the plane. If we write the equation of the plane in intercept form, we can determine the position vector for the x-, y-, and z-intercepts of the plane.
The equation $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$ must be true for all points on a plane. Therefore, we should first rearrange $7x + 3y + z + 12 = 0$ into the form $1 = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$.
$7x + 3y + z = -12$ becomes $\frac{7}{-12}x + \frac{3}{-12}y + \frac{1}{-12}z = 1$
Therefore, $a = \frac{-12}{7}$, $b = \frac{-12}{3} = -4$, and $c = \frac{-12}{1} = -12$ and the position vectors of the three intercepts are $\vec{A} = \langle -1.714, 0, 0 \rangle$, $\vec{B} = \langle 0, -4, 0 \rangle$, and $\vec{C} = \langle 0, 0, -12 \rangle$.
To complete the problem, compute the dot product.
$\left(\vec{B} \times \hat{n}\right)\hat{n} = \left(B_x \hat{n}_x + B_y \hat{n}_y + B_z \hat{n}_z\right)\hat{n} = \left(0\left(\frac{7}{\sqrt{59}}\right) - 4\left(\frac{3}{\sqrt{59}}\right) + 0\left(\frac{1}{\sqrt{59}}\right)\right)\left\langle \frac{7}{\sqrt{59}}, \frac{3}{\sqrt{59}}, \frac{1}{\sqrt{59}} \right\rangle$.
$\left(\vec{B} \times \hat{n}\right)\hat{n} = \frac{-12}{\sqrt{59}}\left\langle \frac{7}{\sqrt{59}}, \frac{3}{\sqrt{59}}, \frac{1}{\sqrt{59}} \right\rangle = \left\langle \frac{-84}{59}, \frac{-36}{59}, \frac{-12}{59} \right\rangle = \langle -1.424, -0.610, -0.203 \rangle$
Useful Links: http://www.math.umn.edu/~nykamp/m2374/readings/lineplane/

6.6 Vector Direction

Learning objectives

- Determine the direction angles which identify the direction of a particular vector.
- Determine the direction cosines which identify the direction of a particular vector.
- Determine the coefficients of a unit vector to represent a particular direction.

Introduction

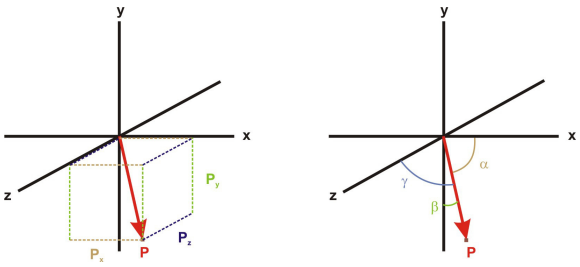
Throughout the previous sections we have used component notation to describe vectors. We have also used unit vectors to indicate a particular direction. In this section we will discuss the use of angles to specify a particular direction.

Direction Angles

As we have seen before, the equation for a vector is given by

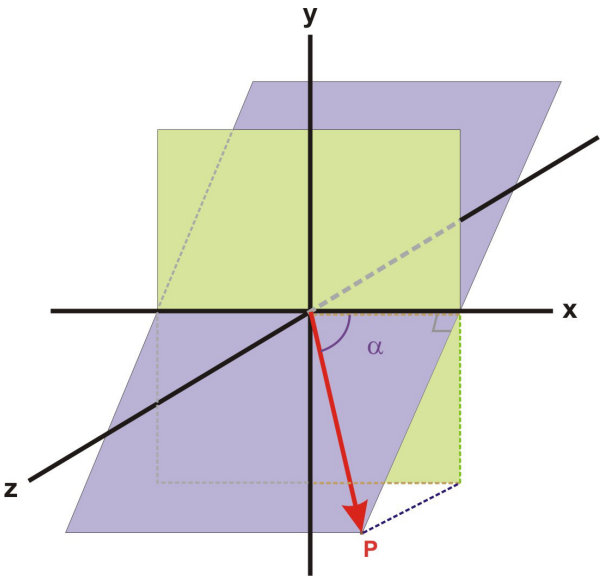
$\vec{p} = \langle P_x, P_y, P_z \rangle$

where P_x, P_y, and P_z are the x, y, and z coordinates of the vector obtained by projecting the vector onto the x, y, and z axes as shown below left.



The image above right, shows the angles between the position vector, $\vec{P}$, and the three axes: is the angle between $\vec{P}$ and the x-axis, is the angle between $\vec{P}$ and the y-axis, and is the angle between $\vec{P}$ and the z-axis.

The position vector, $\vec{P}$, and the unit vector, $\hat{x}$, define a plane shown in lavender below.



The direction angle, , is the angle between $\vec{P}$ and $\hat{x}$ in the plane defined by the two vectors. The other plane shown in the diagram is the X-Y plane, which was included in the diagram to help you visualize orientation of the other plane.

In our discussion of the dot product, we saw that the dot product of two vectors can be given by $\vec{A} \times \vec{B} = |\vec{A}||\vec{B}|\cos\theta$

Therefore, we can calculate the angle between $\vec{P}$ and the unit vector $\hat{x}$.

$\alpha = \cos^{-1}\frac{\vec{P}\times\hat{x}}{|\vec{P}|}$

Similarly, the direction angles and can be calculated using the equations

$\beta = \cos^{-1}\frac{\vec{P}\times\hat{y}}{|\vec{P}|}$ and $\gamma = \cos^{-1}\frac{\vec{P}\times\hat{z}}{|\vec{P}|}$

In some applications, such as astronomy and applied optics, the direction cosines are used at least as frequently as the directional angles themselves.

$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|}, \cos\beta\frac{\vec{P}\times\hat{y}}{|\vec{P}|}$, and $\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|}$

Example: Because airplanes move in three dimensions, air-craft ground crews can use direction cosines to identify their location at any moment. Rather than an arbitrary set of orthogonal x, y, and z axes, the position of an airplane is measured relative to the east, north, and zenith directions. (Zenith means top or upward.) At a particular moment, a small plane is 97 km east, 135 km north, and 7.5 km above its home airport. What are the directional cosines and directional angles for the position vector of the plane at that moment?

Solution: If we use the standard orientation of maps in the northern hemisphere, the x-direction corresponds to east, the y-direction to north, and the z-direction to the zenith. Therefore, the position vector of the plane can be written as $\vec{P} = \langle 297, 135, 7.5\rangle$ km.

The direction cosines associated with this vector are given by

$$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\alpha = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{297}{\sqrt{297^2+135^2+7.5^2}} = 0.582$$

$$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\beta = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{135}{\sqrt{297^2+135^2+7.5^2}} = 0.811$$

$$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\gamma = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{7500}{\sqrt{297^2+135^2+7.5^2}} = 0.045$$

The associated direction angles are

$$= \cos^{-1} 0.582 = 54.4^o$$

$$= \cos^{-1} 0.165 = 35.8^o$$

$$= \cos^{-1} 0.045 = 87.4^o$$

Example: In an upcoming episode of a crime drama, a swarm of special-effect insects bothers one of the investigators after the discovery of a murder victim in a drainage ditch. The animator uses position vectors to track the positions of the virtual pests with respect to an origin at the investigator's head. One such insect is located at a point 33 cm in front, 52 cm to the left, and 18 cm below the tip of the investigator's nose. What are the directional cosines for this insect?

Solution: As we look at our investigator, the +x direction is to her left, the +y direction is upward from her nose, and the +z direction is in front of her. The position vector of the midge can be written as $\vec{P} = \langle 33, 52, -18\rangle$ cm.

The direction cosines associated with this vector are given by

$$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\alpha = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{33}{\sqrt{33^2+52^2+(-18)^2}} = 0.514$$

$$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\beta = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{52}{\sqrt{33^2+52^2+(-18)^2}} = 0.810$$

$$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

$$\cos\gamma = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{-18}{\sqrt{33^2+52^2+(-18)^2}} = 0.281$$

Pythagorean Property of Direction Cosines

An interesting property of direction cosines can be seen if we write the equations for the direction cosines in terms of the components of the position vector, P_x, P_y, and P_z and using the definition of the vector magnitude $|\vec{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2}$

For example,

$$\cos\alpha = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

The other two directional cosines can be rewritten similarly:

$$\cos\beta = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}} \text{ and } \cos\gamma = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}}$$

If we square both sides of all three equations and then sum them, we obtain

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = \frac{P_x^2}{P_x^2+P_y^2+P_z^2} + \frac{P_y^2}{P_x^2+P_y^2+P_z^2} + \frac{P_z^2}{P_x^2+P_y^2+P_z^2}$$

which simplifies to

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = \frac{P_x^2+P_y^2+P_z^2}{P_x^2+P_y^2+P_z^2} = 1$$

This is an important result because the definition of the unit vector stated that $|\hat{u}| = 1$

which also means that

$$u_x{}^2 + u_y{}^2 + u_z{}^2 = \cos^2 + \cos^2 + \cos^2$$

and that the components of the unit vectors correspond to the direction cosines.

Example: A local astronomer used direction cosines when programming the projector in a new planetarium dome. The projector itself sits at the center of the dome, 2.5 m above the floor. He wishes to project Mintaka, one of the stars in the belt of Orion, at a position 12 m south and 2.3 m east of the projector and 8.7 m above the floor. What is the equation of the directional unit vector that the astronomer must enter in to the projection computer?

Solution: We can use the same coordinate system that we used in Example 1 above: $\hat{x}$ = east, $\hat{y}$ = north, and $\hat{z}$ = upward. The projector itself is the origin. In such a coordinate system, the position vector for Mintaka becomes $\vec{P} = \langle 2.3, -12, (8.7-2.5)\rangle = \langle 2.3, -12, 6.2\rangle$. Notice that we did not use Mintaka's position above the floor as the z-coordinate. Rather, since the projector is 2.5 m above the floor, we needed to use the difference between the ceiling height and the projector height. Use the component form of the directional cosine equation to calculate the three components of the unit vector.

$$\cos\alpha = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{2.3}{\sqrt{2.3^2+(-12)^2+6.2^2}} = 0.168$$

$$\cos\beta = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{-12}{\sqrt{2.3^2+(-12)^2+6.2^2}} = -0.876$$

$$\cos\gamma = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}} = \frac{6.2}{\sqrt{2.3^2+(-12)^2+6.2^2}} = 0.453$$

Therefore, the position unit vector is given by $\hat{u} = \langle 0.168, -0.876, 0.453\rangle$

Lesson Summary

The direction of a particular vector can be identified using the components of the vector. The direction could also be defined using the directional angles between the vector and the coordinate axes. The dot product can be used to determine the cosine of the directional angles

$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|}, \cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|}$, and $\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|}$

These angles are also related to the components of the unit vector corresponding to the direction of the vector $\vec{P}$

$u_x{}^2 + u_y{}^2 + u_z{}^2 = \cos^2 + \cos^2 + \cos^2$

Practice Problems

1. Determine the components of the position vector $\vec{P} = \langle 2.4, 5.3, 1.8\rangle$ then determine the directional angles between this vector and the x-axis.
2. Determine the direction cosines for the vector $\vec{N} = \langle 8, 3, -5\rangle$
3. Determine the position vector for a small plane at the moment it is 2.5 km east, 8.8 km south, and 4.1 km above its home airport. Use a coordinate system where the x-direction corresponds to east, the y-direction to north, and the z-direction to the zenith. Then determine the directional cosines used by the air-traffic control personnel to identify the plane's location.
4. Use the method of direction cosines to identify the unit vector having the same direction as the position vector $\vec{R} = \langle 781, 978, 1310\rangle$.
5. Determine the direction angles between the vector $\vec{p} = \langle 25, 8, 15\rangle$ and the coordinate axes.
6. In the crime drama scene described in the example, a swarm of special-effect midges bothers one of the investigators after the discovery of a murder victim in a drainage ditch. One midge is located at a point 7.2 cm in front, 22 cm to the right, and 14 cm above the tip of the investigator's nose. What are the directional cosines used by the special-effects animator to identify this midge?

Solutions

1. Determine the components of the position vector $\vec{P} = \langle 2.4, 5.3, 1.8\rangle$ then determine the directional angles between this vector and the x-axis. $\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$
   $\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{2.4}{\sqrt{(2.4)^2+(5.3)^2+(1.8)^2}} = \frac{2.4}{\sqrt{5.76+28.09+3.24}} = \frac{2.4}{\sqrt{37.09}} = 0.394$
   $\alpha = \cos^{-1}0.394 = 66.8$
2. Determine the direction cosines for the vector $\vec{N} = \langle 8, 3, -5\rangle \cos\alpha = \frac{\vec{N}\times\hat{x}}{|\vec{N}|} = \frac{N_x}{\sqrt{N_x^2+N_y^2+N_z^2}}$
   $\cos\alpha = \frac{\vec{N}\times\hat{x}}{|\vec{N}|} = \frac{N_x}{\sqrt{N_x^2+N_y^2+N_z^2}} = \frac{8}{\sqrt{(8)^2+(3)^2+(-5)^2}} = \frac{8}{\sqrt{64+9+25}} = 0.7213$
   $\cos\beta = \frac{\vec{N}\times\hat{y}}{|\vec{N}|} = \frac{N_y}{\sqrt{N_x^2+N_y^2+N_z^2}}$
   $\cos\beta = \frac{\vec{N}\times\hat{y}}{|\vec{N}|} = \frac{N_y}{\sqrt{N_x^2+N_y^2+N_z^2}} = \frac{3}{\sqrt{(8)^2+(3)^2+(-5)^2}} = \frac{3}{\sqrt{64+9+25}} = 0.3030$
   $\cos\gamma = \frac{\vec{N}\times\hat{z}}{|\vec{N}|} = \frac{N_z}{\sqrt{N_x^2+N_y^2+N_z^2}}$
   $\cos\gamma = \frac{\vec{N}\times\hat{z}}{|\vec{N}|} = \frac{N_z}{\sqrt{N_x^2+N_y^2+N_z^2}} = \frac{-5}{\sqrt{(8)^2+(3)^2+(-5)^2}} = \frac{-5}{\sqrt{64+9+25}} = -0.5051$
3. Determine the position vector for a small plane at the moment it is 2.5 km east, 8.8 km south, and

4.1 km above its home airport. Use a coordinate system where the x-direction corresponds to east, the y-direction to north, and the z-direction to the zenith. Then determine the directional cosines used by the air-traffic control personnel to identify the plane's location. In this coordinate system, with an origin at the home airport, the position vector is given by $\vec{r} = \langle 2.5, -8.8, 4.1 \rangle$ with units of kilometers.

$\cos\alpha = \frac{\vec{r}\times\hat{x}}{|\vec{r}|} = \frac{r_x}{\sqrt{r_x^2+r_y^2+r_z^2}}$

$\cos\alpha = \frac{\vec{r}\times\hat{x}}{|\vec{r}|} = \frac{r_x}{\sqrt{r_x^2+r_y^2+r_z^2}} = \frac{2.5}{\sqrt{(2.5)^2+(-8.8)^2+(4.1)^2}} = \frac{2.5}{\sqrt{100.5}} = 0.249$

$\cos\beta = \frac{\vec{r}\times\hat{y}}{|\vec{r}|} = \frac{r_y}{\sqrt{r_x^2+r_y^2+r_z^2}}$

$\cos\beta = \frac{\vec{r}\times\hat{y}}{|\vec{r}|} = \frac{r_y}{\sqrt{r_x^2+r_y^2+r_z^2}} = \frac{-8.8}{\sqrt{(2.5)^2+(-8.8)^2+(4.1)^2}} = \frac{-8.8}{\sqrt{100.5}} = -0.878$

$\cos\gamma = \frac{\vec{r}\times\hat{z}}{|\vec{r}|} = \frac{r_z}{\sqrt{r_x^2+r_y^2+r_z^2}}$

$\cos\gamma = \frac{\vec{r}\times\hat{z}}{|\vec{r}|} = \frac{r_z}{\sqrt{r_x^2+r_y^2+r_z^2}} = \frac{4.1}{\sqrt{(2.5)^2+(-8.8)^2+(4.1)^2}} = \frac{4.1}{\sqrt{100.5}} = 0.409$

4. Use the method of direction cosines to identify the unit vector having the same direction as the position vector $\vec{R} = \langle 781, 978, 1310 \rangle$. The unit vector which has the same direction as this vector has the components

$\hat{u} = \langle \cos\alpha, \cos\beta, \cos\gamma \rangle$ where $\cos\alpha = \frac{\vec{R}\times\hat{x}}{|\vec{R}|} = \frac{R_x}{\sqrt{R_x^2+R_y^2+R_z^2}}$, $\cos\beta = \frac{\vec{R}\times\hat{y}}{|\vec{R}|} = \frac{R_y}{\sqrt{R_x^2+R_y^2+R_z^2}}$, and $\cos\gamma = \frac{\vec{R}\times\hat{z}}{|\vec{R}|} = \frac{R_z}{\sqrt{R_x^2+R_y^2+R_z^2}}$.

Once we find the three direction cosines, we have the components of the unit vector,

$\cos\alpha = \frac{\vec{R}\times\hat{x}}{|\vec{R}|} = \frac{791}{\sqrt{791^2+978^2+1310^2}} = \frac{791}{\sqrt{329826}} = \frac{791}{1816.11} = 0.436$

$\cos\beta = \frac{\vec{R}\times\hat{y}}{|\vec{R}|} = \frac{978}{\sqrt{791^2+978^2+1310^2}} = \frac{978}{\sqrt{3298265}} = \frac{978}{1816.11} = 0.539$

$\cos\gamma = \frac{\vec{R}\times\hat{z}}{|\vec{R}|} = \frac{1310}{\sqrt{791^2+978^2+1310^2}} = \frac{1310}{\sqrt{3298265}} = \frac{1310}{1816.11} = 0.721$

$\hat{u} = \langle \cos\alpha, \cos\beta, \cos\gamma \rangle = \langle 0.436, 0.539, 0.721 \rangle$

5. Determine the direction angles between the vector $\vec{p} = \langle 25, 8, 15 \rangle$ and the coordinate axes. $\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$

$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{25}{\sqrt{(25)^2+(8)^2+(15)^2}} = \frac{25}{\sqrt{914}} = \frac{25}{30.23} = 0.827$

$\alpha = \cos^{-1}0.827 = 34.2°$

$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}}$

$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{8}{\sqrt{(25)^2+(8)^2+(15)^2}} = \frac{8}{\sqrt{914}} = \frac{8}{30.23} = 0.265$

$\beta = \cos^{-1}0.265 = 75.7°$

$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}}$

$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{15}{\sqrt{(25)^2+(8)^2+(15)^2}} = \frac{15}{\sqrt{914}} = \frac{15}{30.23} = 0.496$

$\gamma = \cos^{-1}\ 0.496 = 60.25°$

6. In the crime drama scene described in the example, a swarm of special-effect midges bothers one of the investigators after the discovery of a murder victim in a drainage ditch. One midge is located at a point 7.2 cm in front, 22 cm to the right, and 14 cm above the tip of the investigator's nose. What are the directional cosines used by the special-effects animator to identify this midge? As we look at our investigator, the +x direction is to her left, the +y direction is upward from her nose, and the +z direction is in front of her. The position vector of the midge can be written as $\vec{P} = \langle -22, 14, 7.2 \rangle$ cm. The direction cosines associated with this vector are given by

$$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{P_x}{\sqrt{P_x^2+P_y^2+P_z^2}}$$
$$\cos\alpha = \frac{\vec{P}\times\hat{x}}{|\vec{P}|} = \frac{-22}{\sqrt{(-22)^2+(14)^2+(7.2)^2}} = \frac{-22}{\sqrt{731.84}} = \frac{-22}{27.05} = -0.813$$
$$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{P_y}{\sqrt{P_x^2+P_y^2+P_z^2}}$$
$$\cos\beta = \frac{\vec{P}\times\hat{y}}{|\vec{P}|} = \frac{14}{\sqrt{(-22)^2+(14)^2+(7.2)^2}} = \frac{14}{\sqrt{731.84}} = \frac{14}{27.05} = 0.518$$
$$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{P_z}{\sqrt{P_x^2+P_y^2+P_z^2}}$$
$$\cos\gamma = \frac{\vec{P}\times\hat{z}}{|\vec{P}|} = \frac{7.2}{\sqrt{(-22)^2+(14)^2+(7.2)^2}} = \frac{7.2}{\sqrt{731.84}} = \frac{7.2}{27.05} = 0.266$$

6.7 Vector Equations

Learning objectives

- Identify a line in space as the intersection of two planes.
- Find points on a line given the vector equation for the line.
- Find the vector equation for a line given two points on the line.

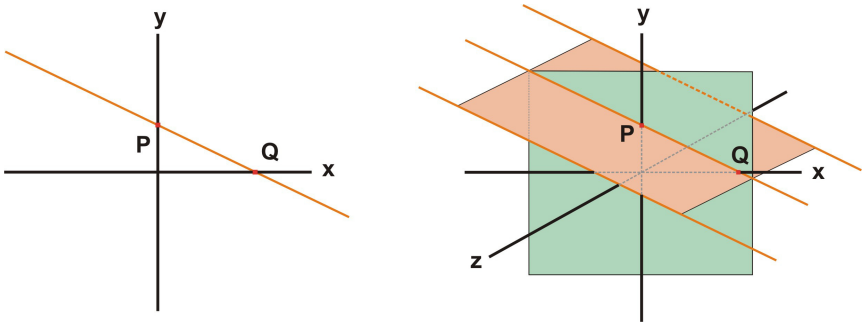
Introduction

You are probably very familiar with using $y = mx + b$, the slope-intercept form, as the equation of a line. While this equation works well in two-dimensional space, it is insufficient to completely define the equation of a line in higher order spaces. To help you see this clearly, several people have developed java-applets which allow one to rotate a vector and its coordinate system. Two of the best such applets can be accessed at http://www.univie.ac.at/future.media/moe/galerie/geom1/geom1.html#gb and http://www.walter-fendt.de/m14e/line3d.htm You may want to spend some time with one or both of these applets before reading the rest of this section.

Lines in Space

In a two-dimensional plane, a line can be represented by the equation $y = mx + b$, where m is the slope of the line and b is the y-intercept. In three-dimensional space, this equation can represent a line or a plane or the projection of a 3D line onto the x-y plane. For example, in the diagram below, the line on the left passes through points P and Q. The equation in 2D space for this line is $y = \frac{-P}{Q}x + P$

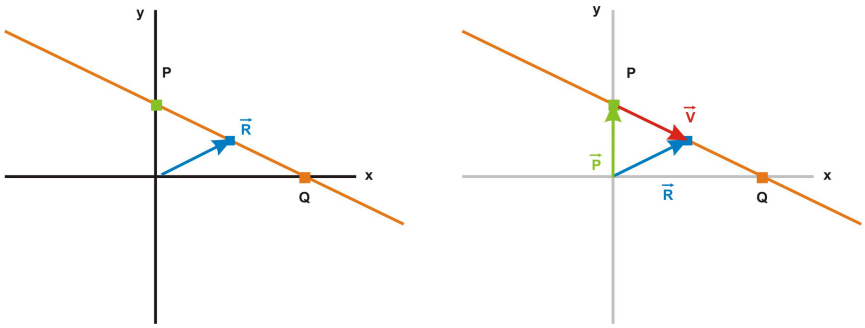
The three orange lines in the diagram at right below can also be described by this equation. All three have slope $M = \frac{-P}{Q}$ and y-intercept $B = P$. If we want to specify one of these lines in particular, we need to identify the line as the intersection of two planes. For example, the line where $y = \frac{-P}{Q}x + P$ and $z = 0$ for all points is the intersection of the orange and green planes in the diagram below right. (Remember that the orange and green planes are drawn with edges for convenience, but that they represent planes that extend outward toward infinity.)



Therefore we can specify a particular line by requiring that the equations for both planes be simultaneously true. In a previous section we developed one equation for a plane given by $d = -n_x\ x_o\ -n_y\ y_o\ -n_z\ z_o$ This means that in this case $Qy + Px - QP = 0$ and $z = 0$ must both be satisfied for all points on the line passing through points P and Q.

Finding the Vector Equation of a Line

Another way to identify points on a line can be found using the vector addition method we discussed earlier in this chapter. To find the position vector, $\vec{r}$, for any point along a line, we can add the position vector of a point on the line which we already know and add to that a vector, $\vec{v}$, that lies on the line as shown in the diagram below.



The position vector $\vec{R}$ for a point between P and Q is given by $\vec{R} = \vec{p} + \vec{v}$

All other points on this line can be reached by traveling along the line from point P, therefore the position vector for any point on the line is given by $\vec{r} = \vec{p} + k\vec{v}$.

If we know the locations of two points on a line, we can determine the equation of the line. All points on a line fulfill the equation $\vec{r} = \vec{p} + k\vec{v}$, where k is a scalar that varies from -∞ to ∞. If we already know the position vectors for two points on the line, $\vec{p}$ and $\vec{q}$, we can use the method of vector subtraction to determine the equation of the vector, $\vec{v} = \vec{p} - \vec{q}$. Therefore, $\vec{r} = \vec{p} + k(\vec{p} - \vec{q})$, where k varies from -∞ to ∞.

Example: Write the equation of $y = -\frac{5}{3}x + 5$ as a vector equation.

Solution: Start by choosing two points on the line, say (3,0) and (6,-5). Then $\vec{p} = \langle 3, 0\rangle$ and $\vec{q} = \langle 6, -5\rangle$. So,

$\vec{p} - \vec{q} = \langle 3, 0\rangle - \langle 6, -5\rangle = \langle 3 - 6, 0 - (-5)\rangle = \langle -3, 5\rangle$

Finally, the vector equation of the line is

$\vec{r} = \langle 3, 0\rangle + k\langle -3, 5\rangle = \langle 3 - 3k, 5k\rangle$

Example: Write the vector equation of the line defined by the points (2, 3, 4) and (-5, 2, -1).

Solution: These two points have position vectors $\langle 2, 3, 4\rangle$ and $\langle -5, 2, -1\rangle$. The vector of the line connecting the two points is given by

$\overrightarrow{v} = \overrightarrow{p} - \overrightarrow{q} = \langle (2-(-5)), (3-2), (4-(-1))\rangle$

$\hat{v} = \overrightarrow{p} - \overrightarrow{q} = \langle 7, -1, 5\rangle$

The equation of the line then becomes

$\overrightarrow{r} = \langle 2, 3, 4\rangle + k\langle 7, -1, 5\rangle \qquad \overrightarrow{r} = \langle 2, 3, 4\rangle + k\langle 7, -1, 5\rangle$

Example 2: Determine the equation for the line defined by the points $P = (6, 7, 5)$ and $Q = (3, 2, 11)$. Then find the position vector for a point, R, half-way between these two points.

Solution:

$\overrightarrow{v} = \overrightarrow{p} - \overrightarrow{q} = \langle (6-3), (7-2), (5-11)\rangle$

$\hat{v} = \overrightarrow{p} - \overrightarrow{q} = \langle 3, 5, -6\rangle$

The equation of the line then becomes

$\overrightarrow{r} = \langle 6, 7, 5\rangle + k\langle 3, 5, -6\rangle \qquad \overrightarrow{r} = \langle 6, 7, 5\rangle + k\langle 3, 5, -6\rangle$

Since the vector $\overrightarrow{v}$ points from P, the value of k for a particular point gives us some information about the location of that point. If $0 < k < 1$, the point lies on the line between P and Q. If $k < 0$, the point is not between P and Q and is closer to P than Q. If $k > 1$, the point lies on the line between P and Q. If $k < 0$, the point is not between P and Q and is closer to P than Q. The point halfway between the two points has $k = \frac{1}{2}$.

$\overrightarrow{R} = \langle 6, 7, 5\rangle + \frac{1}{2}\langle 3, 5, -6\rangle$

$\overrightarrow{R} = \langle (6+1.5), (7+2.5), (5+(-3)\rangle = \langle 7.5, 9.5, 2\rangle$

Example: Air-craft control is tracking two planes in the vicinity of their airport. At a given moment, one plane is at a location 45 km east and 120 km north of the airport at an altitude of 7.5 km. The second plane is located 63 km east and 96 km south of the airport at an altitude of 6.0 km. The first plane is flying directly toward the airport while the second plane is continuing at a constant altitude with a heading defined by the vector $\overrightarrow{h_2} = \langle 3, 4, 0\rangle$ to land eventually at another airport to the northwest of our air-traffic controllers. Do the paths of these two aircraft cross? (Note: Two lines that are not parallel but do not cross are called **skew lines**.)

Solution: The first thing we need to do is to determine the position vectors of the two planes. Define our airport as the origin of coordinates and define $\hat{x}$ = east, $\hat{y}$ = north, and $\hat{z}$ = upward. Call the position of plane #1, P, and the position of plane #2, Q. This gives $\overrightarrow{p} = \langle 45, 120, 7.5\rangle$ and $\overrightarrow{q} = \langle 63, -96, 6.0\rangle$

Plane #1 is heading directly toward the airport. The vector from the position of plane #1 to the origin is given by

$\overrightarrow{v} = \overrightarrow{o} - \overrightarrow{p} = \langle (0-45), (0-120), (0-7.5)\rangle$

$\overrightarrow{v} = \overrightarrow{p} - \overrightarrow{q} = \langle -45, -120, -7.5\rangle$

The equation of the line representing plane #1's motion then becomes

$\overrightarrow{r_1} = \langle 45, 120, 7.5\rangle + k_1\langle -45, -120, -7.5\rangle \qquad \overrightarrow{r_1} = \langle 45, 120, 7.5\rangle + k_1\langle -45, -120, -7.5\rangle$

Plane #2 continues at a constant altitude with a heading $\overrightarrow{h_2} = \langle 3, 4, 0\rangle$. The equation of the line representing plane #2's motion then becomes

$\overrightarrow{r_2} = \langle -63, 96, 6.0\rangle + k_2\langle 3, 4, 0\rangle \qquad \overrightarrow{r_2} = \langle 63, -96, 6.0\rangle + k_2\langle 3, 4, 0\rangle$

$\overrightarrow{r_2} = \langle -63, 96, 6.0\rangle + k_2\langle 3, 4, 0\rangle$

$\overrightarrow{r_1} = \langle 45, 120, 7.5\rangle + k_1 \langle -45, -120, -7.5\rangle$

If the paths of the two planes intersect, the position of that intersection must have coordinates which satisfy both equations.

To satisfy both equations, r_{1x} must equal r_{2x}, r_{1y} must equal r_{2y}, and r_{1z} must equal r_{2z}. This means that 45 - $45k_1$ = 63 + $3k_2$, 120 - $120k_1$ = -96 + $4k_2$, and 7.5 - $7.5k_1$ = 6.0 + $0k_2$ must all be simultaneously true.

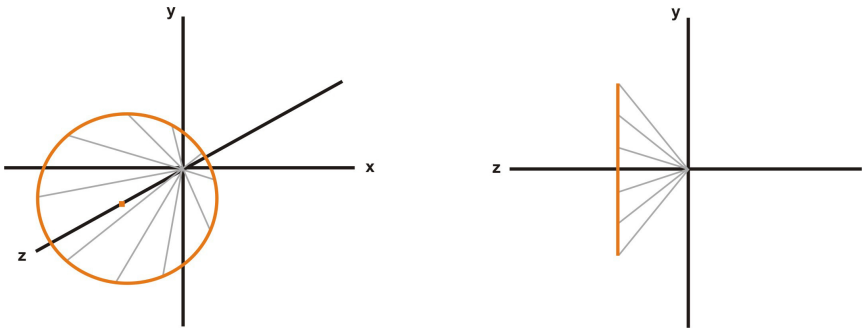
The first equation simplifies to 15 - $15k_1$ = 21 + k_2 or k_2 = -6 - $15k_1$. The second simplifies to 30 - $30k_1$ = -24 + k_2 or k_2 = 54 - $30k_1$. The third equation simplifies to 7.5 - $7.5k_1$ = 6 or $k_1 = \frac{1}{5}$.

Inserting the value of k_1 from the third equation into the first two equations, we find $k_2 = -6 - 15\left(\frac{1}{5}\right) = -6 - 3 = -9$ and $k_2 = 54 - 30k_1 = 54 - 30\left(\frac{1}{5}\right) = 54 - 6 = 48$ Since $k_2 = 33$ and $k_2 = 0$ cannot be simultaneously true, the two planes' paths do not cross.

Note: There are other ways to reason through this problem. Can you find some? (Hint: Think of the geometry in 2-dimensions, or focus on the altitude as a way to simplify this problem).

General Vector Equations

In general, a vector equation is any function that takes any one or more variables and returns a vector. The vector equation of a line is an equation that identifies the position vector of every point along the line. This works for straight lines and for curves. For example, the vector equation $\overrightarrow{p} = \langle 3 \cos\theta, 3 \sin\theta, 2\rangle$ defines a circle having a radius of 3 units which sits parallel to the x-y plane at a distance of 2 units along the z-axis as shown below. The orange line is the set of points that satisfy the vector equation $\overrightarrow{p} = \langle 3 \cos\theta, 3 \sin\theta, 2\rangle$ while the grey lines are various individual position vectors which satisfy the vector equation.



Lesson Summary

In two-dimensional space, a line is defined by the equation $y = mx + b$ where m is the slope of the line and b is the y-intercept. In three-dimensional space, a line is defined by the intersection of two planes. In other words, the line is the set of points for which the equations of the two planes are simultaneously true. Also, if we know the position vector $\overrightarrow{p}$ of one point on the line and the unit vector $\hat{u}$ in the direction of the line, all other points on the line satisfy the equation $\overrightarrow{r} = \overrightarrow{p} + k\hat{u}$. In general, a vector equation is any function that takes any one or more variables and returns a vector.

Practice Problems

1. Write the vector equation of the line defined by the points (5, -7, 3) and (2, 6, -4).

2. Are the two vectors $\vec{D} = \langle 1, -1, 4\rangle + d\langle 1, -1, 1\rangle$ and $\vec{F} = \langle 2, 4, 7\rangle + f\langle 2, 1, 3\rangle$ skew lines or do they intersect one another?
3. In physics, the motion of an object traveling at a constant speed is described by the equation $\vec{s_t} = \vec{s_i} + \vec{v}t$ where s_i is the initial position, s is the position at some later time t, and v is the velocity of the object. Write the vector equation which returns the set of position vectors $\vec{s}$ for an object having an initial position $\vec{s_i} = \langle 2, 3, 4\rangle$ and a velocity of $\vec{v} = \langle 1, 1, -2\rangle$ and determine the object's location at t = 10s.
4. An object has a position of $\vec{s_i} = \langle 3, 3, 6\rangle$ at $t = 0$ and a velocity of $\vec{v} = \langle 10, 7, 3\rangle$. Use the vector equation $\vec{s} = \vec{s_i} + \vec{v}t$ to determine the distance traveled by the object between $t = 3$s and $t = 5$s. Distance measured in meters.
5. Determine the vector equation of the straight line defined by the points (2, 2, 2) and (1, 3, 5).
6. Do the two lines $\vec{R} = \langle 4, 4, -2\rangle + r\langle -3, 7, 2\rangle$ and $\vec{K} = \langle 9, -8, 7\rangle + k\langle -2, 1, 3\rangle$ intersect?

Solutions

1. Write the vector equation of the line defined by the points (5, -7, 3) and (2, 6, -4). These two points have position vectors $\langle 5, -7, 3\rangle$ and $\langle 2, 6, -4\rangle$. The vector of the line connecting the two points is given by
   $\vec{v} = \vec{p} - \vec{q} = \langle (5-2), (-7-6), (3-(-4))\rangle = \langle 3, -13, 7\rangle$
   The equation of the line, $\vec{r} = \vec{q} + k\vec{v}$, then becomes
   $\vec{r} = \langle 2, 6, -4\rangle + k\langle 3, -13, 7\rangle \qquad \vec{r} = \langle 5, -7, 3\rangle + k\langle 3, -13, 7\rangle$
2. Are the two vectors $\vec{D} = \langle 1, -1, 4\rangle + d\langle 1, -1, 1\rangle$ and $\vec{F} = \langle 2, 4, 7\rangle + f\langle 2, 1, 3\rangle$ skew lines or do they intersect one another? If the two vectors intersect, there must be a point identified by position vector $\vec{p}$ which satisfies the equations of both lines. In other words, we must be able to find values for d and f such that $\vec{D} = \vec{F}$ or $\langle 1, -1, 4\rangle + d\langle 1, -1, 1\rangle = \langle 2, 4, 7\rangle + f\langle 2, 1, 3\rangle$.
   Each of the three components of vectors $\vec{D}$ and $\vec{F}$ must independently be equal if $\vec{D} = \vec{F}$. This means that $1 + d = 2 + 2f$, -1 - d = $4 + f$, and $4 + d = 7 + 3f$ Combining the first two equations gives: $1 + d$ - 1 - $d = 2 + 2f + 4 + f$ which when simplified becomes $0 = 6 + 3f$ or f = -2. If we substitute this back into the first equation we obtain, $1 + d = 2 + 2$(-2) which when simplified becomes d = -3. If the two lines cross, f = -2 and d = -3 should satisfy each of the component equations.
   $1 + d = 2 + 2f$ becomes 1 + (-3) = 2 + 2(-2) or -2 = -2
   -1 - d = $4 + f$ becomes -1 - (-3) = 4 + (-2) or +2 = +2
   $4 + d = 7 + 3f$ becomes 4 + (-3) = 7 + 3(-2) or +1 = +1
   All three equations are equally satisfied, so the two lines do intersect and the point of intersection is (-2, 2, 1).
3. In physics, the motion of an object traveling at a constant speed is described by the equation $\vec{s_t} = \vec{s_i} + \vec{v}t$ where s_i is the initial position, s is the position at some later time t, and v is the velocity of the object. Write the vector equation which returns the set of position vectors $\vec{s}$ for an object having an initial position $\vec{s_i} = \langle 2, 3, 4\rangle$ and a velocity of $\vec{v} = \langle 1, 1, -2\rangle$ and determine the object's location at t = 10s.$\vec{s_t} = \vec{s_i} + \vec{v}t = \langle 2, 3, 4\rangle + t\langle 1, 1, -2\rangle = \langle 1 + t, 3 + t, 4 - 2t\rangle$
   At t = 10s, $\vec{s_{10}} = \langle 1 + 10, 3 + 10, 4 - 2(10)\rangle = \langle 11, 13, -16\rangle$
4. An object has a position of $\vec{s_i} = \langle 3, 3, 6\rangle$ at $t = 0$ and a velocity of $\vec{v} = \langle 10, 7, 3\rangle$. Use the vector equation $\vec{s} = \vec{s_i} + \vec{v}t$ to determine the distance traveled by the object between $t = 3$s and $t = 5$s. Distance measured in meters.The vector equation describing the motion of the object is
   $\vec{s} = \vec{s_i} + \vec{v}t = \langle 3, 3, 6\rangle + t\langle 10, 7, 3\rangle = \langle 3 + 10t, 3 + 7t, 6 + 3t\rangle$
   The object's position at $t = 3$s is obtained from the vector equation:
   $\vec{s_3} = \vec{s_i} + \vec{v}t = \langle 3, 3, 6\rangle + (3)\langle 10, 7, 3\rangle = \langle 3 + 10(3), 3 + 7(3), 6 + 3(3)\rangle = \langle 33, 24, 15\rangle$
   The object's position at $t = 5$s is obtained from the vector equation:
   $\vec{s_5} = \vec{s_i} + \vec{v}t = \langle 3, 3, 6\rangle + (5)\langle 10, 7, 3\rangle = \langle 3 + 10(5), 3 + 7(5), 6 + 3(5)\rangle = \langle 53, 38, 21\rangle$

The distance traveled between these two points is the magnitude of the vector starting at (33, 24, 15) and ending at (53, 38, 21).
$\overrightarrow{\triangle s} = \overrightarrow{s_5} - \overrightarrow{s_3} = \langle 53, 38, 21\rangle - \langle 33, 24, 15\rangle = \langle 20, 14, 6\rangle$
Now we can use the Pythagorean theorem to determine the magnitude of the vector.
$|\overrightarrow{\triangle s}| = \sqrt{a^2 + b^2 + c^2} = \sqrt{(20)^2 + (14)^2 + (6)^2} = 25.1$ meters

5. Determine the vector equation of the straight line defined by the points (2, 2, 2) and (1, 3, 5). These two points have position vectors $\overrightarrow{p} = \langle 2, 2, 2\rangle$ and $\overrightarrow{q} = \langle 1, 3, 5\rangle$. The vector of the line connecting the two points is given by
$\overrightarrow{v} = \overrightarrow{p} - \overrightarrow{q} = \langle (2-1), (2-3), (2-5)\rangle = \langle 1, -1, -3\rangle$
The equation of the line, $\overrightarrow{r} = \overrightarrow{p} + k\overrightarrow{v}$, then becomes
$\overrightarrow{r} = \langle 2, 2, 2\rangle + k\langle 1, -1, -3\rangle$ $\overrightarrow{r} = \langle 2, 2, 2\rangle + k\langle 1, -1, -3\rangle$
6. Do the two lines $\overrightarrow{R} = \langle 4, 4, -2\rangle + r\langle -3, 7, 2\rangle$ and $\overrightarrow{K} = \langle 9, -8, 7\rangle + k\langle -2, 1, 3\rangle$ intersect? If the two vectors intersect, there must be a point identified by position vector $\overrightarrow{p}$ which satisfies the equations of both lines. In other words, we must be able to find values for r and k such that $\overrightarrow{R} = \overrightarrow{K}$ or $\langle 4, 4, -2\rangle + r\langle -3, 7, 2\rangle = \langle 9, -8, 7\rangle + k\langle -2, 1, 3\rangle$.
Each of the three components of vectors $\overrightarrow{R}$ and $\overrightarrow{K}$ must independently be equal if $\overrightarrow{R} = \overrightarrow{K}$. This means that $4 - 3r = 9 - 2k$, $4 + 7r = -8 + k$, and $-2 + 2r = 7 + 3k$
Solving the second equation for k gives $12 + 7r = k$. Now substitute this into one of the other two equations: $4 - 3r = 9 - 2(12 + 7r)$ or $4 - 3r = 9 - 24 - 14r$ which simplifies to $19 = -11r$ or $r = \frac{-19}{11} = -1.727$. Substitute this value into one of the other equations: $-2 + 2(-1.727) = 7 + 3k$ which becomes $-2 - 3.454 = 7 + 3k$. Solving for k gives $k = -4.151$. If the two lines cross, $r = -1.727$ and $k = -4.151$ should satisfy each of the component equations.
$4 - 3r = 9 - 2k$ becomes $4 - 3(-1.727) = 9 - 2(-4.151)$ or $-1.181 = 17.302$ Since even this first equation does not hold true, these two lines are skew and do not intersect.

6.8 Applications of Vector Analysis

Learning objectives

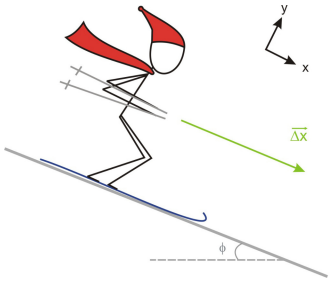
- Use the dot product to determine the work done on a moving object by a given force.
- Use the cross product to determine the force exerted on a moving charge by a magnetic field.
- Use the cross product to determine the torque exerted by a force on a object.

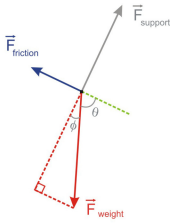
Introduction

Vectors can be used by air-traffic controllers when tracking planes, by meteorologists when describing wind conditions, and by computer programmers when they are designing virtual worlds. In this section, we will present three applications of vectors which are commonly used in the study of physics: work, torque, and magnetic force.

Work

In physics, the term **work** is used to describe energy that is added to or removed from an object or system when a force is applied to it. From experiment, it has been determined that work is maximized when the applied force is parallel to the motion of the object and that no work is done when the force is applied perpendicular to the motion. Therefore, the work done by a force can be described by the dot product of the force vector and the displacement vector. For example, several forces act on the skier in the diagram below.





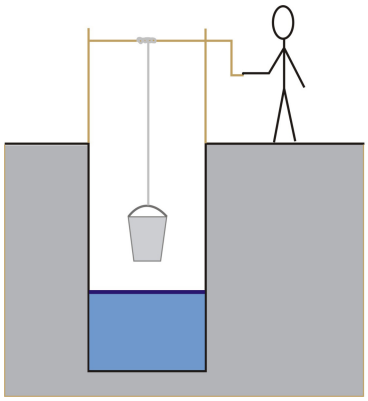
Sven's weight pulls downward toward the center of the earth, the snow support's Sven by pushing upward on his skis perpendicular to the slope, and the friction between Sven's skis and the snow points in the opposite direction from his motion. The work done by each of these forces can be determined using the dot product of the force and the displacement vector $\overrightarrow{\triangle x}, W = \overrightarrow{F} \times \overrightarrow{\triangle x} = F(\triangle x) \cos \theta$.

$W_{friction} = \overrightarrow{F_{friction}} \times \overrightarrow{\triangle x} = (F_{friction})(\triangle x) \cos 180^\circ = -F_{friction} \triangle x$

$W_{support} = \overrightarrow{F_{support}} \times \overrightarrow{\triangle x} = (F_{support})(\triangle x) \cos 90^\circ = 0$

$W_{weight} = \overrightarrow{F_{weight}} \times \overrightarrow{\triangle x} = (F_{weight})(\triangle x) \cos \theta = (F_{weight})(\triangle x) \sin \phi$

Example: Xiao turns a crank to lower a bucket of water into a well. Determine the total work done on the bucket if the weight of the bucket is 15 N and the tension force in the rope is 13 N. The bucket rises a distance of 4.5 m while he is cranking.



Solution: The work done by a force acting on an object is described by the dot product of the force vector and the displacement vector: $W = \overrightarrow{F} \times \overrightarrow{\triangle x} = F(\triangle x) \cos \theta$. In this case, the rope does negative work on the bucket because the motion and the force are in opposite directions. If the force is measured in newtons and the displacement in meters, the work is measured in Joules.

$W_{rope} = \overrightarrow{F_{rope}} \times \overrightarrow{\triangle x} = F_{rope}(\triangle x) \cos \theta = (13N)(4.5m) \cos 180^\circ = -58.5J$

The weight force does positive work on the bucket because the motion and the force are in the same direction.

$W_{weight} = \overrightarrow{F_{weight}} \times \overrightarrow{\triangle x} = F_{weight}(\triangle x) \cos \theta = (15N)(4.5m) \cos 0^\circ = 67.5J$

The total work is the sum of the two individual amounts of work.

$W_{total} = W_{rope} + W_{weight} = -58.5 + 67.5J = 9.0J$

A total of 9.0 J of work is done on the bucket as the bucket moves downward into the well.

Magnetic Force

The force that a magnetic field exerts on a charged particle is strongest when the particle moves perpendicular to the field and the magnetic force on the particle is equal to zero when it moves parallel to the field. Therefore the magnetic force can be described using the cross-product of the field strength vector and the particle's velocity vector: $\vec{F} = q\vec{v} \times \vec{B}$ where $\vec{F}$ is the force on the particle, q is the charge of the particle, $\vec{v}$ is the velocity of the particle, and $\vec{B}$ is the vector representing the magnetic field. If the velocity is measured in m/s and if the magnetic field is measured in tesla, the force will be measured in newtons, the metric base-unit of force.

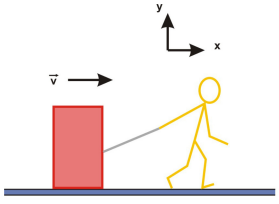
Example: In a cathode-ray-tube EKG monitor, the heart-beat trace on the screen is drawn by a beam of electrons that move from the back of the tube toward the screen. When the electrons hit the screen, the collision transfers energy from the electrons to the phosphor on the inner surface of the screen, which then glows. The point where the electrons hit the screen can be changed by changing the magnetic field inside the tube. A sensor connected to the patient translates the electrical pulses across the heart into the strength of the magnetic field inside the EKG's cathode ray tube, which then changes the path of the electrons. The force which the magnetic field exerts on the electrons is determined both by the strength of the magnetic field and on the velocity of the electrons. If the electrons move at 25,000 m/s northward through a 5.2 millitesla magnetic field which points west to east across the tube, what is the magnitude and direction of the force exerted on the electron. The electron has a charge of $q_{electron} = -1.6 \times 10^{-19} C$.

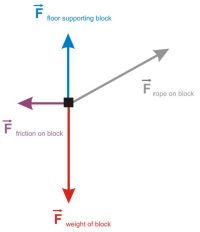
Torque

When you lift a baseball off a table-top, you are exerting a force that moves the object as a whole. When you apply a force to a doorknob, you cause the door to rotate on its hinges. Scientists use the term **torque** to describe the force-like property that affects the rotation of an object. The torque can be described using the cross-product of the force vector and the lever arm, a vector pointing radially outward from the axis of rotation to the point where the force is applied to the object: $\vec{\tau} = \vec{r} \times \vec{F}$, where $\vec{\tau}$ is the torque, $\vec{r}$ is the lever arm, and $\vec{F}$ is the applied force.

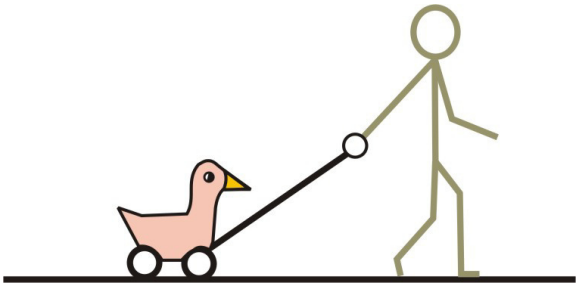
Practice Problems

1. The work done by a force on an object is given by the dot product of the force vector and the displacement vector representing the motion of the object. The diagram below shows Sanjay pulling a large crate across the floor. The four forces which act on the crate during this process are shown in the diagram below. Which of the four forces exert non-zero force on the crate?

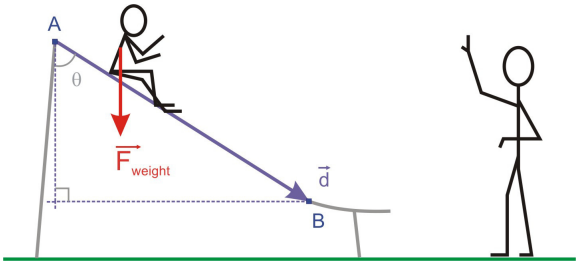




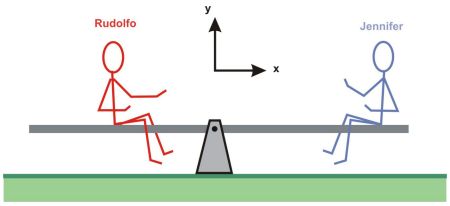
2. Le'andra is pulling her toy duck (mass 0.75 kg) at a constant speed of 3.0 m/s. The string she uses to pull the duck makes an angle of 42° above the horizontal and Le'andra keeps a constant tension in the string of 2.0 N. What is the amount of work done by the tension force when the duck is pulled forward a distance of 2.8 m?

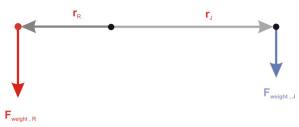


3. Beuford has once again taken Brynna to the park to play on the slide. If Brynna has a weight of 25 kg and if the slide has a 30° incline above the horizontal, what work is done by her weight as she slides down the 3.5 m incline? Remember that the weight force in newtons is equal to the product of the mass and the acceleration of gravity, 9.8 m/s^2.



4. The scientists at Fermi Lab in Chicago, IL use magnetic fields to direct beams of protons during their explorations of the submicroscopic structure of atoms and quarks. Both magnetic field strength and the velocity of the protons are vector quantities. Determine the force on a proton moving northward at 4.2×10^6 m/s through a magnetic field of 2.5 T oriented from east to west. (Note: protons are very tiny and are therefore able to move VERY fast.)
5. Vector $\vec{B}$ represents the magnitude and direction of the magnetic field in a certain region. Vector $\vec{v}$ represents the velocity of a charged particle, $q = 3.2 \times 10^{-15}$ C, which moves through the magnetic field. The lengths of the vectors are such that the velocity vector is measured in m/s and the magnetic field vector is measured in tesla. What is the force experienced by the charged particle as it moves through the magnetic field?
6. If two children, Rudolfo (weight = 210 N) and Jennifer (175 N), sit on either end of a see-saw as shown below. Rudolfo is 1.0 m from the pivot and Jennifer is 1.4 m from the pivot. What torques are exerted by the children on the see-saw?

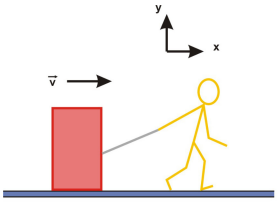


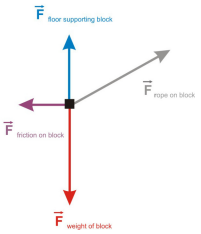


7. Exercise scientists and physical therapists use torque to analyze various exercises such as the triceps exercise shown in the diagram below. Here, the triceps muscle exerts a force on the elbow-end of the forearm, affecting the rotation of the forearm.
   If the triceps exerts a force of 17N, what torque is applied to the forearm?

Solutions

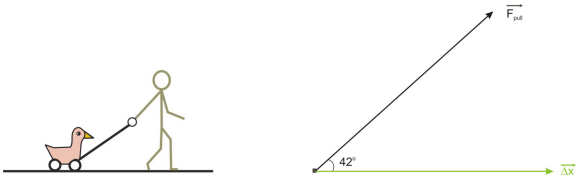
1. The diagram below shows Sanjay pulling a large crate across the floor. The four forces which act on the crate during this process are shown in the diagram below. Which of the four forces exert non-zero force on the crate?





The dot product is defined by $\vec{A} \times \vec{B} = |A||B| \cos\theta$ therefore only forces which have at least some component parallel to the motion will do non-zero work on the object. The angle between the displacement and forces perpendicular to the motion is 90° so $\vec{A} \times \vec{B} = |A||B| \cos\theta = 0$. The force from the floor and the weight of the crate do no work, since both of these forces are perpendicular to the motion of the crate. The rope does positive work on the crate since the force of the rope on the crate has a non-zero x-component. The friction does negative work on the crate since it is in the opposite direction from the displacement.

2. De'andra is pulling her toy duck (mass 0.75 kg) at a constant speed of 3.0 m/s. The string she uses to pull the duck makes an angle of 42° above the horizontal and De'andra keeps a constant tension in the string of 2.0 N. What is the amount of work done by the tension force when the duck is pulled forward a distance of 2.8 m?

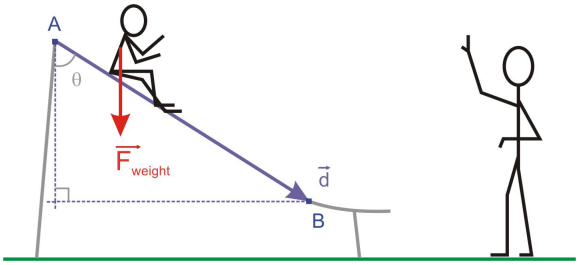


The work done by De'andra on the duck depends on the force she uses to pull the duck and on the distance the duck moves while she pulls. It also depends on the angle between the pulling force and the displacement vector.

$\overrightarrow{F_{pull}} \times \vec{x} = |\overrightarrow{F_{pull}}||\vec{x}| \cos\theta = (2.0N)(2.8m) \cos 42° = 4.16Nm$

The N represents "newton", the unit of force. The m represents "meters," the unit of displacement. (1.0 N)(1.0 m) = 1.0 J where J represents "joules," the unit of work and energy.

3. Beuford has once again taken Brynna to the park to play on the slide. If Brynna has a weight of 25 kg and if the slide has a 30° incline above the horizontal, what work is done by her weight as she slides down the 3.5 m incline? Remember that the weight force in newtons is equal to the product of the mass and the acceleration of gravity, 9.8 m/s^2.



If the slide is inclined at 30° above the horizontal, then = 60° from the vertical. The work done by a force on an object is given by the dot product of the force and the displacement of the object. Here $\overrightarrow{F_{weight}} = mg = (25kg)(9.8m/s^2) = 245N$. Therefore,

$W = \overrightarrow{F_{weight}} \times \vec{d} = |\overrightarrow{F_{weight}}||\vec{d}| \cos\theta = (245N)(3.5m) \cos 60° = 428.75J$

4. The scientists at Fermi Lab in Chicago, IL use magnetic fields to direct beams of protons during their explorations of the submicroscopic structure of atoms and quarks. Both magnetic field strength and the velocity of the protons are vector quantities. Determine the force on a proton moving northward

at 4.2×10^6 m/s through a magnetic field of 2.5 T oriented from east to west. (Note: protons are very tiny and are therefore able to move VERY fast.) Define a coordinate system where eastward is the +z direction, northward is the +y direction, and upward is the +z direction. In this coordinate system, $\vec{v} = \langle 0, 4.2 \times 10^6, 0 \rangle$ and $\vec{B} = \langle -2.5, 0, 0 \rangle$. Therefore, $|\vec{F}| = q|\vec{v} \times \vec{B}| = q\left(|\vec{v}||\vec{B}| \sin\theta\right)$. Since the velocity is northward and the magnetic field is westward, the angle between the two vectors is 90^o.

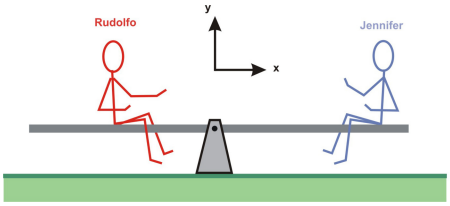
$|\vec{F}| = q|\vec{v} \times \vec{B}| = q\left(|\vec{v}||\vec{B}| \sin\theta\right) = (1.6 \times 10^{-19})\left((4.2 \times 10^6)(2.5) \sin 90^\circ\right) = 16.8 \times 10^{-13} N$

Using the right hand rule, we can determine the direction of the force on the proton. If you point your thumb northward along the velocity vector and your fore-finger westward along the magnetic field vector, your palm and your extended middle-finger point upward. Therefore the force which the magnetic field exerts on the proton is in the +z direction: $|\vec{F}| = q\vec{v} \times \vec{B} = \langle 0, 0, 16.8 \times 10^{-13} N \rangle$.

5. Vector $\vec{B}$ represents the magnitude and direction of the magnetic field in a certain region. Vector $\vec{v}$ represents the velocity of a charged particle, $q = 3.2 \times 10^{-15}$ C, which moves through the magnetic field. The lengths of the vectors are such that the velocity vector is measured in m/s and the magnetic field vector is measured in tesla. What is the force, $\vec{F}$, experienced by the charged particle as it moves through the magnetic field?We saw in a previous problem that $|\vec{F}| = q|\vec{v} \times \vec{B}|$, therefore can use the component version of the cross product equation to solve this problem.

$\vec{F} = q\left(\vec{v} \times \vec{B}\right) = q\left\langle (v_y B_z - v_z B_y), (v_z B_x - v_x B_z), (v_x B_y - v_y B_x) \right\rangle$

6. If two children, Rudolfo (weight = 210 N) and Jennifer (weight = 175 N), sit on either end of a see-saw as shown below. Rudolfo is 1.0 m from the pivot and Jennifer is 1.4 m from the pivot. What torques are exerted by the children on the see-saw?



The coordinate system has been defined such that the weight-force vectors are parallel to the y-axis and the lever-arm vectors are parallel to the x-axis. First determine the component form of each vector equation and then use the component version of the cross-product equation to determine the torque exerted by each child.

$\vec{\tau_R} = \vec{r_R} \times \overrightarrow{F_{weight,R}} = \langle -1.0m, 0, 0 \rangle \times \langle 0, -210N, 0 \rangle$

$\vec{r} \times \vec{F} = \left\langle (r_y F_z - r_z F_y), (r_z F_x - r_x F_z), (r_x F_y - r_y F_x) \right\rangle$

$\vec{r_R} \times \vec{F_R} = \langle ((0 * 0) - (0 * -210)), ((0 * 0) - (-1 * 0)), ((-1 * -210) - (0 * 0)) \rangle = \langle 0, 0, 210mN \rangle$

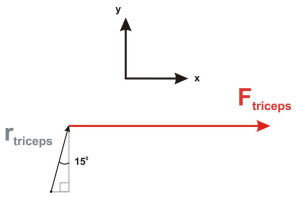
$\vec{\tau_J} = \vec{r_J} \times \overrightarrow{F_{weight,J}} = \langle 1.4m, 0, 0 \rangle \times \langle 0, -170N, 0 \rangle$

$\vec{r} \times \vec{F} = \left\langle (r_y F_z - r_z F_y), (r_z F_x - r_x F_z), (r_x F_y - r_y F_x) \right\rangle$

$\vec{r_J} \times \vec{F_J} = \langle ((0 * 0) - (0 * -170)), ((0 * 0) - (1.4 * 0)), ((1.4 * -170) - (0 * 0)) \rangle = \langle 0, 0, -238mN \rangle$

7. Exercise scientists and physical therapists use torque to analyze various exercises such as the triceps exercise shown in the diagram below. Here, the triceps muscle exerts a force on the elbow-end of the forearm, affecting the rotation of the forearm. If the triceps exerts a force of 17N, what torque is applied to the forearm?

A close-up of the triceps force and lever-arm is shown below.



Since the forearm is positioned at an angle of 15° to the vertical, the angle between the two vectors is 90° – 15° = 75°. The magnitude of the lever-arm vector is the distance from the elbow-pivot to the point where the triceps pulls on the bone, $|\vec{r}| = 2.5cm$. Similarly, the magnitude of the force vector is the strength of the force, $|\vec{F}| = 17N$. Since we know the magnitudes of both vectors and the angle between them, we can use the angle-version of the cross-product equation to determine the magnitude of the torque.

$|\vec{\tau}| = rF \sin\theta = (2.5cm)(17N) \sin 75 = 41.05cm \cdot N$

Chapter 7

Introduction to Calculus

7.1 Limits (An Intuitive Approach)

Learning objectives

- Demonstrate an understanding of the relationship between the tangent line and the limit.
- Demonstrate an understanding of the area between a curve and the x-axis.
- Understand the notion of the limit of a function as an independent variable approaches a finite or infinite value.

Introduction

The discovery of calculus was motivated by two fundamental geometric problems: finding the tangent line to a curve and finding the area of a planar region. In this section, we will show that these two problems are related to a deeper concept of calculus known as the *limit* of a function.

Table 7.1:

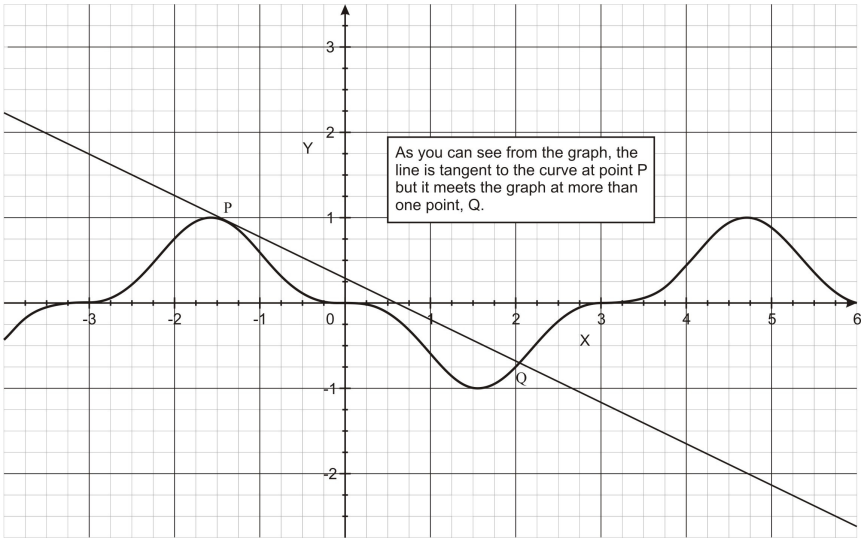
The Two Fundamental Problems of Calculus that Lead to its Discovery:
1. **The Tangent Line Problem:** What is the equation of the tangent line to the graph $f(x)$ at point $P(x_0, y_0)$?
2. **The Area Problem:** What is the area under the graph $f(x)$ and the x-axis in the interval $[a, b]$?

The portion of calculus that deals with the tangent problem is called *differential calculus* and the portion that deals with the area problem is called *integral calculus.* In order to solve those two problems, we need to have a more precise understanding of what a tangent line is and what is meant by the area under a curve. Both of these issues require us to understand a deeper concept, the *limit* of a function.

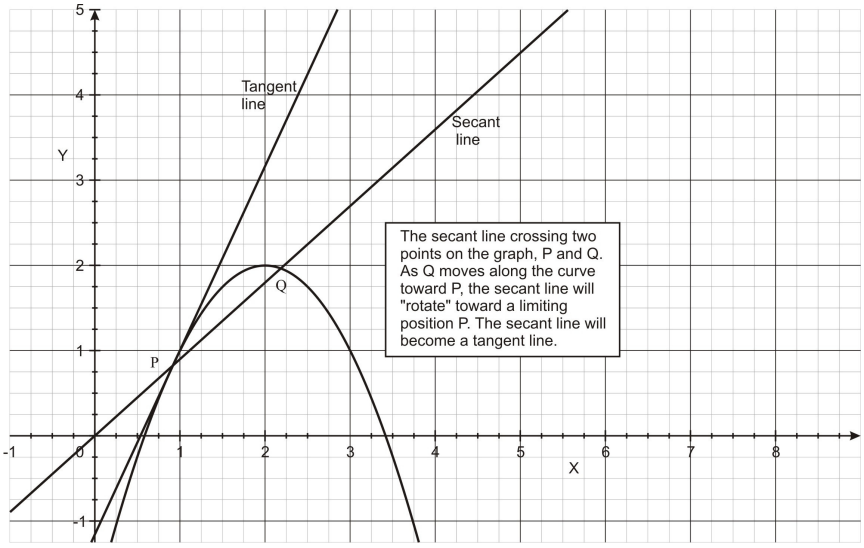
Tangent Lines and Limits

From your studies in geometry, you know that the tangent line is a line that intersects the circle at one point. However, this definition is not precise when we try to apply it to other kinds of curves. For example,

as Figure 1 shows, one can draw a tangent line to a curve yet it cuts the curve at more than one point.



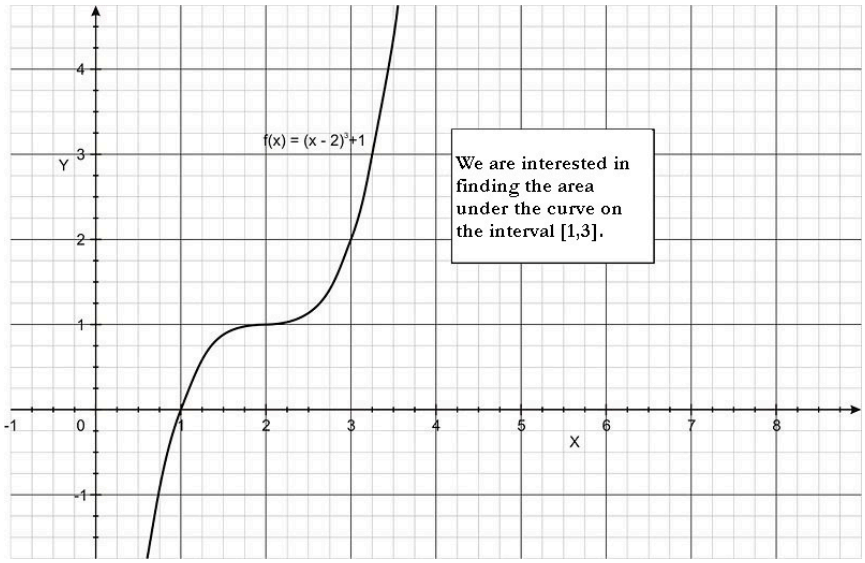
So we need to renew our concept of the tangent line and extend it to apply to curves other than circles. To do so, consider point P on the curve in the figure below. If point Q is any other point on the curve that is different from P, the line that passes through P and Q is called the *secant line*. Imagine if we move point Q along the curve toward point P, the secant line in this case will "rotate" toward a *limiting* position at point P. Eventually, the secant line will become a tangent line at point P, as the figure below shows. This is a new concept of the tangent line, where the general notion of a tangent line leads to the concept of limit. We will deal with the tangent line in more detail in lesson 8.3.

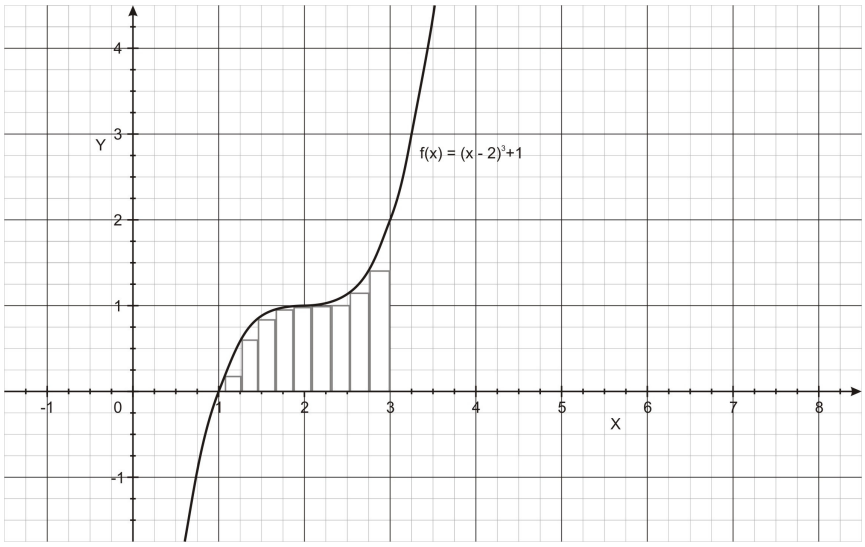

Tangent line
Secant line
The secant line crossing two points on the graph, P and Q. As Q moves along the curve toward P, the secant line will "rotate" toward a limiting position P. The secant line will become a tangent line.
P
Q
X
Y

Area as a limit

Suppose we are interested in finding the area under the curve of a function on the interval $[a, b]$. For example, consider function $f(x) = (x - 2)^3 + 1$ (*Figure a*). Let's say we want to *approximate* the area under the curve from $x = 1$ to $x = 3$. One way to do it is to inscribe rectangles of equal widths on the interval $[1, 3]$ under the curve and then add the areas of these rectangles (*Figure b*). Intuition tells us that if we repeat the process using more and more rectangles to fill the gaps under the curve, our approximation will approach the exact value of the area under the curve. So, the *limiting* value of this approximation is the

exact value of the area under the curve. If we denote the width of each rectangle by Δx and the value of the area under the curve by A, then as Δx approaches zero (the widths of the rectangles get thinner and thinner, and thus less and less gaps), then the area A under the curve will reach an exact value.





What we have seen so far is that the concepts of tangent line and area rest on the notion of *limit*. In the next sections, we will explore those concepts in more details and show how the limit can help us calculate the rate of change of a given quantity. First, however, we introduce some useful notations.

Table 7.2:

Definition of a Limit (an informal view)
The notation $\lim_{x \to x_0} f(x) = L$ means that as x approaches (or gets very close to) x_0, the limit of the function $f(x)$ gets very close to the value L. We first used this notation in Chapter 1.

Example 1:

Make a conjecture about the value of the limit of $\lim_{x\to 0} \frac{3x}{\sqrt{x+1}-1}$.

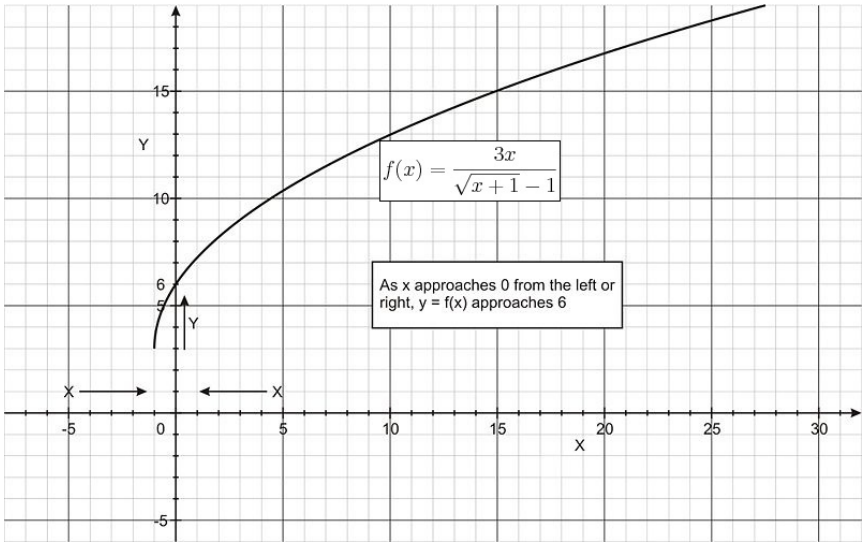
Solution: Notice that the function $f(x) = \frac{3x}{\sqrt{x+1}-1}$ is not defined at $x = 0$. The table below shows samples of x-values approaching 0 from the left side and from the right side. In both cases, the values of $f(x)$, calculated to at least 5 decimal places, get closer and closer to 6. Thus our conjecture is that $\lim_{x\to 0} \frac{3x}{\sqrt{x+1}-1} = 6$.

Table 7.3:

x								
0	0.00001	0.0001	0.001	0.01				
$f(x)$	5.984962	5.9985	5.99985	5.999985	*Undefined*	6.000015	6.00015	6.0015	6.014963

Another way of seeing this is to graph $f(x)$ (shown below). Notice that the x-values approach 0 from the left side and from the right side. In both cases, the values of $f(x)$ appear to get closer and closer to 6.

Hence, again our conjecture is that $\lim_{x\to 0} \frac{3x}{\sqrt{x+1}-1} = 6$.



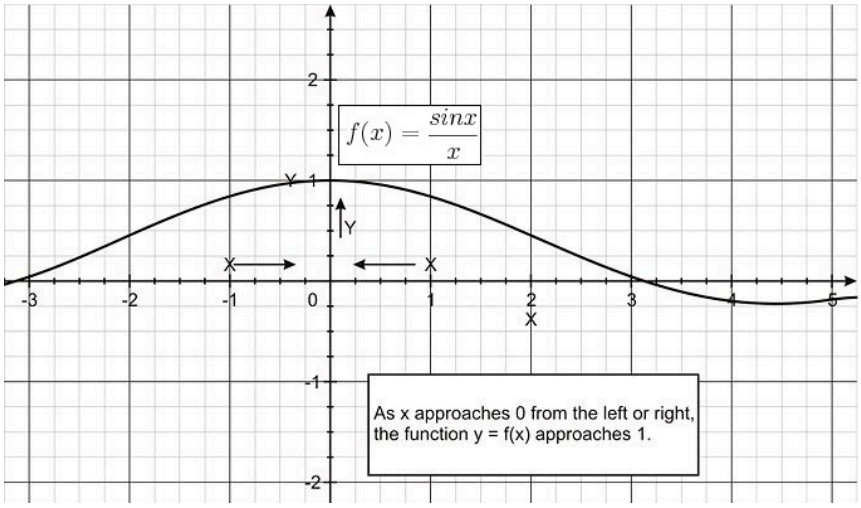
Example 2: Make a conjecture about the value of the limit $\lim_{x\to 0} \frac{sinx}{x}$ where x is in radians.

Solution: The function here is not defined at $x = 0$. With the help of a computing utility, we can obtain the table below.

Table 7.4:

x							
0	0.01	0.1	0.2				
$f(x)$	0.993347	0.998334	0.999983	*Undefined*	0.999983	0.998334	0.993347

The data in the table suggest thats $\lim_{x\to 0} \frac{sinx}{x} = 1$. The graph below supports this hypothesis.



The One-Sided Limits

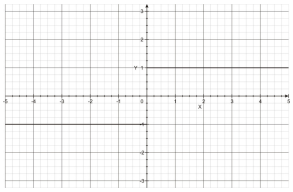
The limit in the definition above and from Examples 1 and 2,

$$\lim_{x \to x_o} f(x) = L$$

is called a two-sided limit because it requires $f(x)$ to get closer and closer to L from both sides of $x = x_0$.However, not all functions behave this way. Some actually have different limits on the two sides of $x = x_0$. For example, consider the function

$$f(x) = \frac{|x|}{x} = \begin{cases} 1, x > 0 \\ -1, x < 0 \end{cases}$$

which is shown in the graph below:



Note that as x approaches 0 from the right, $f(x)$ approaches 1. On the other hand, as x approaches 0 from the left, the function $f(x)$ approaches -1. Mathematically, we write

$$\lim_{x \to 0^+} \frac{|x|}{x} = 1$$
$$\lim_{x \to 0^-} \frac{|x|}{x} = -1$$

Where the superscript "+" indicates a limit from the right and the superscript "-" indicates a limit from the left.

Table 7.5:

The One-Sided Limit

If $f(x)$ approaches L as x approaches x_0 from the left and from the right, then we write

$\lim_{x \to x_0^+} f(x) = L$
$\lim_{x \to x_0^-} f(x) = L$

which reads: "the limit of $f(x)$ as x approaches x_0^+ (or x_0^-) from the right (or left) is L.

Sometimes the value of $f(x)$ does not get closer and closer to some *single* value L as $x \to x_0$. In this case we say that the limit as x approaches x_0 *does not exist.* For example, the two-sided limit $lim_{x \to 0} \frac{|x|}{x}$ that we have just encountered does not exist because the values of $f(x)$ do not approach a single number as x approaches 0. Actually, the values approach -1 from the left and 1 from the right. This leads to an important condition that must be satisfied in order for the limit of a function to exist. It is shown in the box below.

Table 7.6:

Conditions For a Limit to Exist (The relationship between one-sided and two-sided limits)

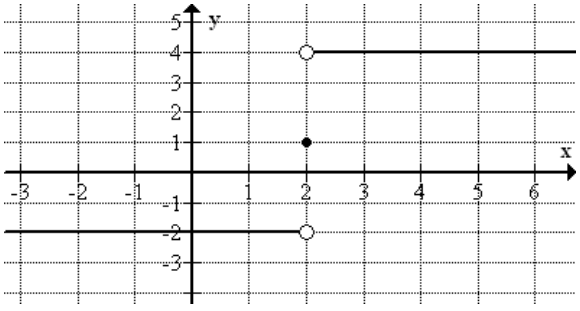
In order for the limit L of a function to exist, both of the one-sided limits must exist at x_0 and must have the *same* value. Mathematically,

$\lim_{x \to x_0} f(x) = L$ if and only if $\lim_{x \to x_0^-} f(x) = L$ and $\lim_{x \to x_0^-} f(x) = L$.

Example 3:

Consider the function f graphed in the accompanying figure and find

1. $\lim_{x \to 2^-} f(x)$
2. $\lim_{x \to 2^+} f(x)$
3. $\lim_{x \to 2} f(x)$
4. $f(2)$



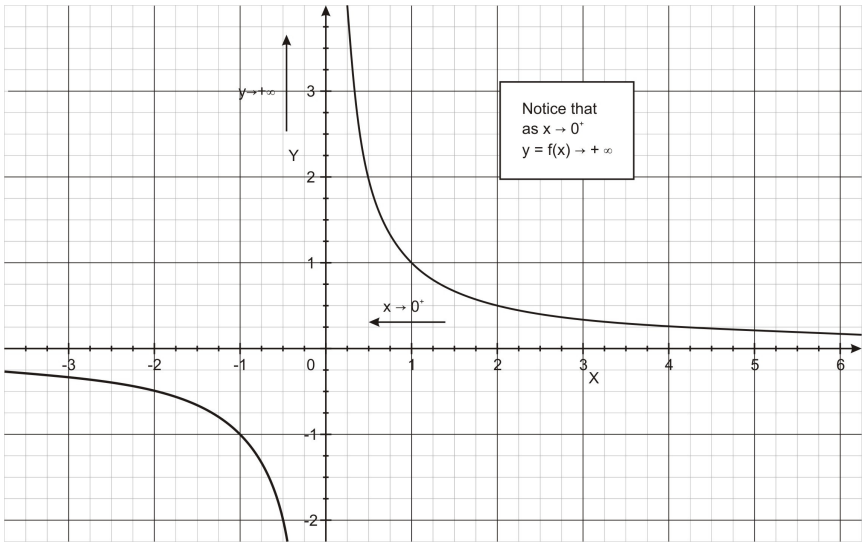
Solution:

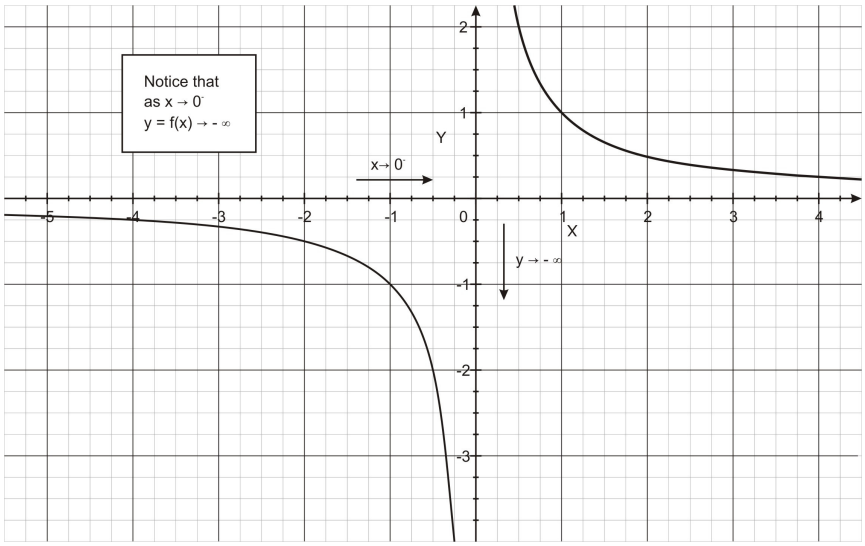
1. From graph, we can see that, $\lim_{x \to 2^-} f(x) = -2$.

2. We can also see from the graph that $\lim_{x \to 2^+} f(x) = 4$.
3. Since the limits from the right and the left are not equal (they do not approach a single value L), the limit does not exist. That is, $\lim_{x \to 2} f(x)$ does not exist.
4. $f(2) = 1$.

Infinite Limits

As previously discussed, the limit of a function may not exist. For example, for the function $f(x) = 1/\text{x}$ (shown in the figures below), as x values are taken closer and closer to 0 from the right, the function increases indefinitely. Also, as x values are taken closer and closer to 0 from the left, the function decreases indefinitely.





We describe these limiting behaviors by writing

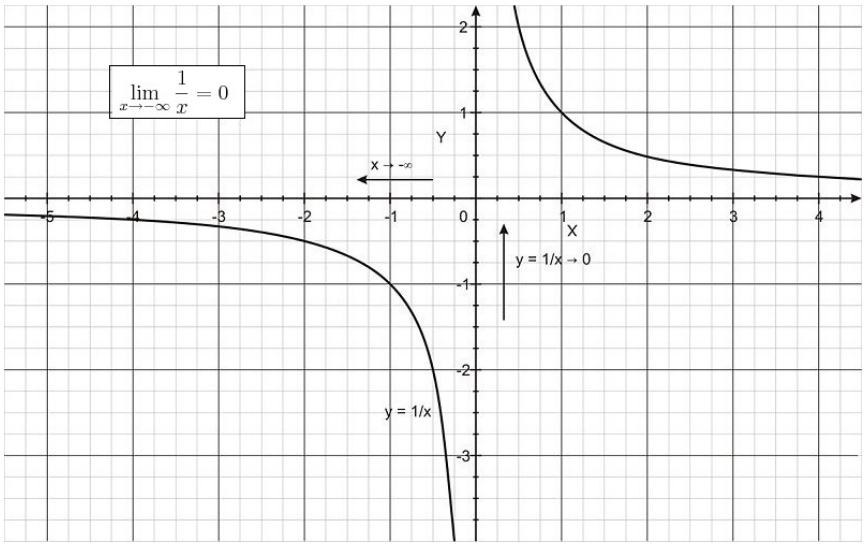
$$\lim_{x \to 0^+} \frac{1}{x} = +\infty$$
$$\lim_{x \to 0^-} \frac{1}{x} = -\infty$$

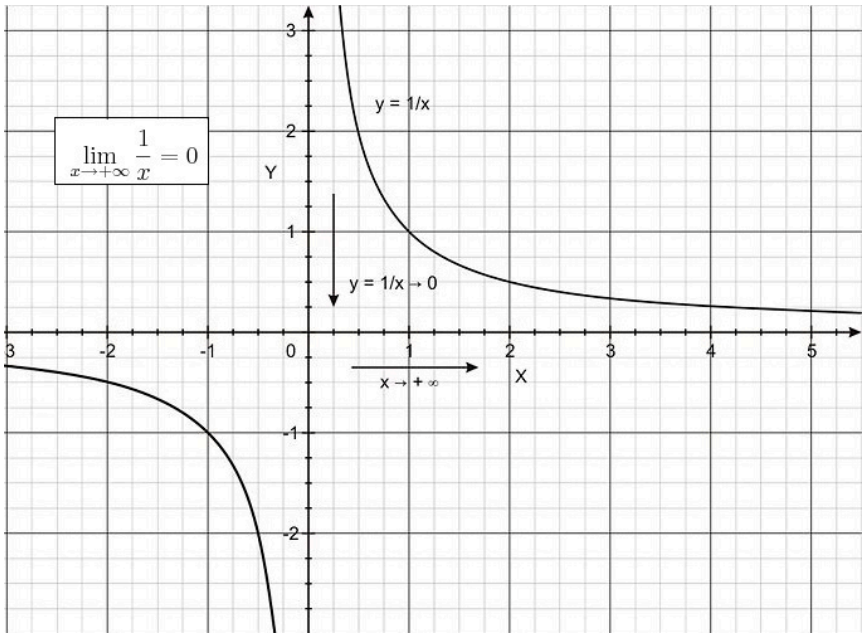
Sometimes we want to know the behavior of $f(x)$ as x increases or decreases without bound. In this case

we are interested in the end behavior of the function, a concept that was first discussed in Chapter 1. For example, what is the value of $f(x) = 1/\text{x}$ as x increases or decreases without bound? That is,

$\lim_{x\to+\infty} \frac{1}{x} = ?$
$\lim_{x\to-\infty} \frac{1}{x} = ?$

As you can see from the graphs (shown below), as x decreases without bound, the values of $f(x) = 1/\text{x}$ are negative and get closer and closer to 0. On the other hand, as x increases without bound, the values of $f(x) = 1/\text{x}$ are positive and still get closer and closer to 0.



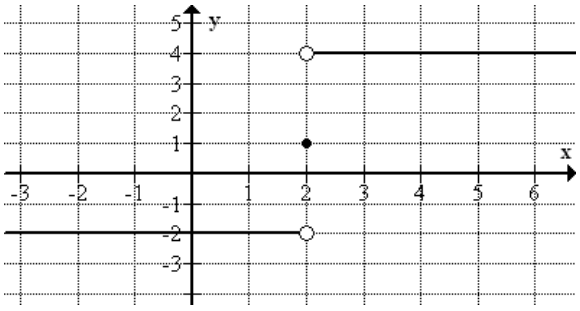


That is,

$\lim_{x\to+\infty} \frac{1}{x} = 0$
$\lim_{x\to-\infty} \frac{1}{x} = 0$

Example 4: Use the graph from Example 3 (shown again below) and find the limits

1. $\lim_{x \to +\infty} f(x)$
2. $\lim_{x \to -\infty} f(x)$



Solution: As you can see from the graph,

1. $\lim_{x \to +\infty} f(x) = 4$, in other words, as x becomes very large, $f(x)$ stays at the same value, 4.
2. $\lim_{x \to +\infty} f(x) = -2$, which says that as x becomes smaller and smaller, $f(x)$ stays at the same value, -2.

Lesson Summary

In this lesson we have taken a first look at the fundamental problems of calculus, the tangent line problem and the area problem, by investigating the concept of a limit. While in chapter 1 we focused on limits of functions in the context of their end behavior, in this lesson we started to find the limit of a function as it approaches any point by using tables and graphs. We learned about one-sided limits and discovered that for the limit of a function to exist as it approaches a given point, both one-sided limits of the function must exist as it approaches the same point. In the next lesson we will continue our study of limits and learn more techniques for computing limits.

Points to Consider

1. What is the benefit of knowing a function's one-sided limits as it approaches a given point?
2. What are examples of functions that have limits that do not exist? Do you see connections between these functions?

Review Questions

1. Make a conjecture about the value of the limit $\lim_{x \to 0} \frac{1-cosx}{x^2}$ where x is in radians. (*Hint:* Make a table of x - values approaching x from the right and the left.)
2. Make a conjecture about the value of the limit $\lim_{x \to 1} \frac{lnx}{2x-2}$. (*Hint:* Make a table of x-values approaching x from the right and the left.)
3. Make a conjecture about the value of the limit $\lim_{x \to 0} \frac{tan2x}{x}$ where x is in radians. (*Hint:* Make a table of x - values approaching x from the right and the left.)
4. Make a graph to evaluate the limit $\lim_{x \to \infty} \frac{x+2}{x+3}$.
5. From Problem 4, evaluate $\lim_{x \to 0} \frac{x+2}{x+3}$.
6. Make a graph to evaluate the limit $\lim_{x \to \infty} \frac{1}{\sqrt{x}}$.
7. From Problem 6, evaluate $\lim_{x \to 0^+} \frac{1}{\sqrt{x}}$
8. Let

Table 7.7:

$f(x) = \begin{cases} x-1, & x \le 3 \\ 3x-7, & x > 3 \end{cases}$

9.
10. and find a.
11. $\lim_{x \to 3^-} f(x)$
12. b.
13. $\lim_{x \to 3^+} f(x)$
14. c.
15. $\lim_{x \to 3} f(x)$
16. Does $\lim_{x \to 0} sin\left(\frac{1}{x}\right)$ exist? Explain why or why not.
17. Use a calculator (or, better, a spread sheet) to find the exact value of $\lim_{x \to 0}(1+x)^{1/x}$

Review Answers

1. 1/2
2. 1/2
3. 2
4. 1
5. 2/3
6. 0
7. $+\infty$
8. a. 2 b. 2
   c. 2
9. *Hint:* It is best to graph it first. You will notice that the graph oscillates between two numbers (what are they?) as it approaches zero.
10. e

7.2 Computing Limits

Learning objectives

- Demonstrate an understanding of the limit of a function from an algebraic perspective.
- Learn how to compute the limit for different kinds of functions.

Introduction

In this lesson, we will present the limit of a function from an algebraic perspective. We will base the algebraic technique on the results obtained from the previous section. We first present the basic theorems that provide the tools necessary to calculate limits. These theorems are outlined in the two boxes below.

Box #1: Important Theorems of Limits

Let a be a real number and suppose that $\lim_{x \to a} f(x) = L_1$ and $\lim_{x \to a} g(x) = L_2$.

Then:

1. $\lim_{x\to a}[f(x)+g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = L_1 + L_2$, meaning the limit of the sum is the sum of the limits.

2. $\lim_{x\to a}[f(x)-g(x)] = \lim_{x\to a} f(x) - \lim_{x\to a} g(x) = L_1 - L_2$, meaning the limit of the difference is the difference of the limits

3. $\lim_{x\to a}[f(x)g(x)] = (\lim_{x\to a} f(x))(\lim_{x\to a} g(x)) = L_1 L_2$, meaning the limit of the product is the product of the limits.

4. $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \frac{L_1}{L_2} L_2 \neq 0$, meaning the limit of a quotient is the quotient of the limits (provided that the denominator does not equal zero.)

5. $\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x\to a} f(x)} = \sqrt[n]{L_1} L_1 > 0$ if n is even, meaning the limit of the n^{th} root is the n^{th} of the limit.

Other useful results follow from the above theorems:

Box #2

1. If a and k are real numbers, then $\lim_{x\to a} k = k$. That is, if $f(x) = k$, a constant function, then the values of $f(x)$ do not change as x is varied, thus the limit of $f(x)$ is k.

2. If a is a real number then $\lim_{x\to a} x = a$. That is, since $f(x) = x$ is an identity function (its input equals its output), then as $x \to a$, $f(x) = x \to a$.

3. $\lim_{x\to a}(k \cdot f(x)) = (\lim_{x\to a} k) \cdot (\lim_{x\to a} f(x)) = k \cdot (\lim_{x\to a} f(x))$

4. $\lim_{x\to a} x^n = (\lim_{x\to a} x)^n = a^n$

Limits of Polynomial Functions

Example 1: Find $\lim_{x\to 1}(x^2 - 3x + 4)$ and justify each step.

Solution: Using Equation (1) of Box #1 (the limit of the sum is the sum of the limits) we get

Table 7.8:

$\lim_{x\to 1}(x^2 - 3x + 4) = \lim_{x\to 1} x^2 + \lim_{x\to 1}(-3x) + \lim_{x\to 1} 4$

From Equation (4) of Box #2, the first term becomes

Table 7.9:

$\lim_{x\to 1} x^2 = (1)^2 = 1$

From Equations (2) and (3) of Box #2, the second term becomes

Table 7.10:

$\lim_{x\to 1}(-3x) = -3 \lim_{x\to 1} x = (-3)(1) = -3$

Finally, from Equation (1) of Box #2, the third term becomes

Table 7.11:

$\lim_{x\to 1} 4 = 4$

Thus the limit of the above polynomial is

Table 7.12:

$\lim_{x\to 1}(x^2 - 3x + 4) = 1 + (-3) + 4 = 2$

However, for conciseness, we can compute the limit by simply substituting $x = 1$ directly into the polynomial,

Table 7.13:

$\lim_{x\to 1}(x^2 - 3x + 4) = (1)^2 - 3(1) + 4 = 2$

and get the same answer. This leads us to a general theorem about limits of polynomials.

Table 7.14:

Theorem: The limit of a polynomial
For any polynomial $f(x) = \mathrm{c_n x^n} + \ldots + \mathrm{c_1 x} + \mathrm{c_0}$ and any real number a, $\lim_{x\to a} f(x) = c_n(a)^n + ... + c_1(a) + c_0$ $\lim_{x\to a} f(x) = f(a)$ In other words, the limit of the polynomial is simply equal to $f(a)$.

Example 2: Find $\lim_{x\to 3}(4x^3 - 4x - 5)$

Solution: According to the theorem above, $\lim_{x\to 3}(4x^3 - 4x - 5) = 4(3)^3 - 4(3) - 5 = 91$

Example 3: Find $\lim_{x\to 5} \frac{x^2-4}{3x^2-2}$

Solution: Using Equation (4) of Box #1 (the limit of the quotient is the quotient of the limit),

Table 7.15:

$\lim_{x\to 5} \frac{x^2-4}{3x^2-2} = \frac{\lim_{x\to 5}(x^2-4)}{\lim_{x\to 5}(3x^2-2)}$

Making use of the limit of the polynomials theorem, we obtain,

Table 7.16:

$\lim_{x\to 5} \frac{x^2-4}{3x^2-2}$	$= \frac{(5)^2-4}{3(5)^2-2}$
	$=\frac{21}{73}$

Limits of Rational Functions

In Example 3 we found the limit of a rational function. Sometimes finding the limit of a rational function at a point a is difficult because evaluating the function at the point a leads to a denominator equal to zero. The box below describes finding the limit of a rational function.

Table 7.17:

Theorem: The limit of a Rational Function

For the rational function $f(x) = \frac{p(x)}{q(x)}$ and any real number a,

$\lim_{x\to a} f(x) = \frac{p(a)}{q(a)}$ if $q(a) \neq 0$.

However, if $q(a) = 0$ then the function may or may not exist. See Examples 5, 6, and 7 below.

Example 4: Find $\lim_{x\to 3} \frac{2-x}{x-2}$.

Solution: Using the theorem above, we simply substitute $x = 3$: $\lim_{x\to 3} \frac{2-x}{x-2} = \frac{2-3}{3-2} = -1$

Example 5: Find $\lim_{x\to 3} \frac{x+1}{x-3}$.

Solution: Notice that the domain of the function is continuous (defined) at all real numbers except at $x = 3$. If we check the one-sided limits we see that $\lim_{x\to 3^+} \frac{x+1}{x-3} = \infty$ and $\lim_{x\to 3^-} \frac{x+1}{x-3} = -\infty$. Because the one-sided limits are not equal, the limit does not exist.

Example 6: Find $\lim_{x\to 2} \frac{x^2-4}{x-2}$.

Solution: Notice that the function here is discontinuous at $x = 2$, that is, the denominator is zero at $x = 2$. However, it is possible to remove this discontinuity by canceling the factor x-2 from both the numerator and the denominator and then taking the limit:

Table 7.18:

$\lim_{x\to 2} \frac{x^2-4}{x-2} = \lim_{x\to 2} \frac{(x-2)(x+2)}{x-2} = \lim_{x\to 2}(x+2) = 4$

This is a common technique used to find the limits of rational functions that are discontinuous at some points. When finding the limit of a rational function, always check to see if the function can be simplified. We will use this technique again in the next example.

Example 7: Find $\lim_{x\to 3} \frac{2x-6}{x^2+x-12}$.

Solution: The numerator and the denominator are both equal to zero at $x = 3$, but there is a common factor x - 3 that can be removed (that is, we can simplify the rational function):

Table 7.19:

$\lim_{x\to 3} \frac{2x-6}{x^2+x-12}$	$= \lim_{x\to 3} \frac{2(x-3)}{(x+4)(x-3)}$
	$= \lim_{x\to 3} \frac{2}{x+4}$
	$= \frac{2}{7}$

Computing Limits Using One-Sided Limits

When we wish to find the limit of a function $f(x)$ as it approaches a point a and we cannot evaluate $f(x)$ at a because it is undefined at that point, we can compute the function's one-sided limits in order to find the desired limit. If its one-sided limits are the same, then the desired limit exists and is the value of the one-sided limits. If its one-sided limits are not the same, then the desired limit does not exist. This technique is used in the examples below.

Example 8: Find the limit $f(x)$ as x approaches 1. That is, find $\lim_{x\to 1} f(x)$ if

Table 7.20:

$f(x) = \begin{cases} 3-x, & x<1 \\ 3x-x^2, & x>1 \end{cases}$

Solution: Remember that we are not concerned about finding the value of $f(x)$ at x but rather *near* x. So, for $x < 1$ (limit from the left),

Table 7.21:

$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-}(3-x) = (3-1) = 2$

and for $x > 1$ (limit from the right),

Table 7.22:

$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+}(3x-x^2) = 2$

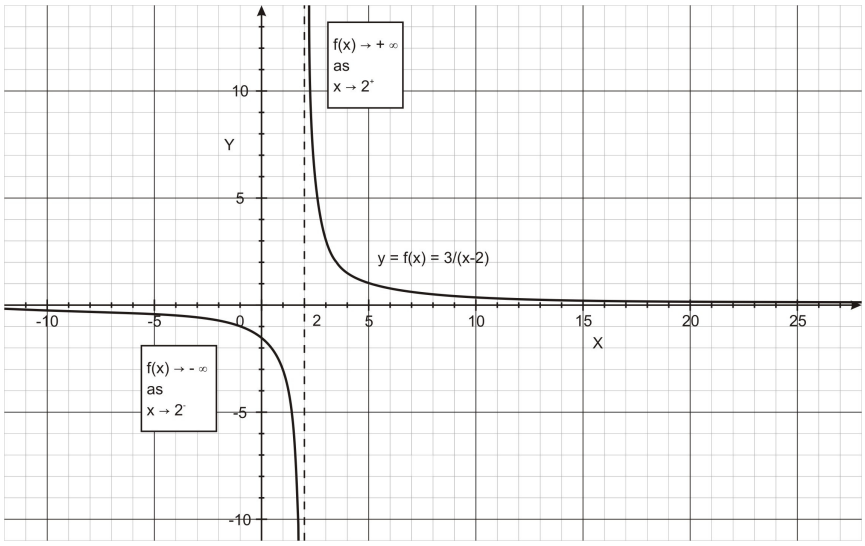
Now since the limit exists and is the same on both sides, it follows that

Table 7.23:

$\lim_{x\to 1} f(x) = 2$

Example 9: Find $\lim_{x\to 2} \frac{3}{x-2}$.

Solution: From the figure below we see that $f(x) = \frac{3}{x-2}$ decreases without bound as x approaches 2 from the left and $f(x) = \frac{3}{x-2}$ increases without bound as x approaches 2 from the right.



This means that $\lim_{x\to 2^-} \frac{3}{x-2} = -\infty$ and $\lim_{x\to 2^+} \frac{3}{x-2} = +\infty$. Since $f(x)$ is unbounded (infinite) in either directions, the limit does not exist.

Lesson Summary

In this lesson we learned many techniques for computing limits. In particular, we learned that we can find the limit of a polynomial function approaching a point a by evaluating the function at a. We also learned that the limits of some rational functions do not exist, but sometimes we can simplify the functions in order to compute their limits. Just because a function is discontinuous at a given point does not necessarily mean that it's limit does not exist at that point. Finally, we learned that limits can often be computed by finding both one-sided limits and checking to see if they are equal.

Points to Consider

1. Can you prove that the theorems in the boxes at the beginning of this lesson are true?
2. Why does it make sense that a limit can exist for a function at a point of discontinuity (as in Example 6)?

Review Questions

1. Find $\lim_{x\to -2}(2)$.
2. Find $\lim_{x\to 0^+}(\pi)$.
3. Find $\lim_{x\to 2} \frac{x^2-4}{x-2}$.
4. Find $\lim_{x\to 6} \frac{x-6}{x^2-36}$.
5. Find $\lim_{x\to 5} \sqrt{x^3-2x-1}$.
6. Find $\lim_{x\to 3^+} \frac{3}{x-3}$.
7. Show that $\lim_{x\to 0^+} \left(\frac{1}{x} - \frac{1}{x^2}\right) = -\infty$.
8. For an object in free fall, such as a stone falling off a cliff, the distance $y(t)$ (in meters) that the object falls in t seconds is given by the kinematic equation $y(t) = 4.9\ t^2$. The object's velocity after 2 seconds is given by $v(t) = \lim_{t\to 2} \frac{y(t)-y(2)}{t-2}$. What is the velocity of the object after 2 seconds?

Review Answers

1. 2
2. π
3. 4
4. $\frac{1}{12}$
5. $\sqrt{114}$
6. $+\infty$
7. Hint: Use a graph.
8. 19.6 m/sec

7.3 Tangent Lines and Rates of Change

Learning objectives

- Demonstrate an understanding of the rate of change of a function.
- Demonstrate an understanding of the slope of the tangent line to the graph.
- Demonstrate an understanding of the instantaneous rate of change.

Introduction

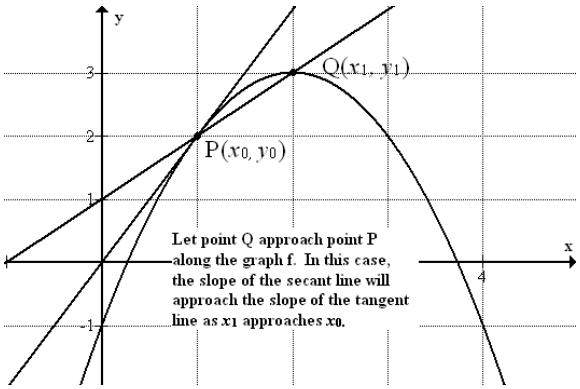
Calculus is a branch of mathematics that deals with rates of change of a quantity. For example, a car speeding down the street, the inflation of currency, the number of bacteria in a culture, and the AC voltage of an electric signal are all examples of quantities that change with time. In this section, we will study the rate of change of a quantity and how it is related to the tangent lines on a curve.

The Tangent Line

Back in Lesson 8.1, we briefly discussed how a secant line could become a tangent line to a point on a graph. Recall from algebra, if points $P(x_0,y_0)$ and $Q(x_1,y_1)$ are two different points on the curve $y = f(x)$, then the *slope* of the secant line connecting the two points is given by

Table 7.24:

m_{sec}	$= \frac{y_1-y_0}{x_1-x_0} = \frac{f(x_1)-f(x_0)}{x_1-x_0}$	(1)


y
x
3
2
1
-1
4
Q(x1, y1)
P(x0, y0)
Let point Q approach point P along the graph f. In this case, the slope of the secant line will approach the slope of the tangent line as x1 approaches x0.

As discussed earlier, if we let the point x_1 approach x_o then Q will approach P along the graph f and thus the slope of the secant line will gradually approach the slope of the tangent line as x_1 approaches x_0. Therefore, (1) becomes

Table 7.25:

	m_{sec}	$= \lim_{x_1 \to x_0} \frac{f(x_1)-f(x_0)}{x_1-x_0}$	(2)

To simplify our notation, if we let $h = x_1 - x_0$, then $x_1 = x_0 + h$ and $x_1 \to x_0$ becomes equivalent to $h \to 0$. This means that (2) becomes

Table 7.26:

	m_{sec}	$= \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$

Table 7.27:

The Slope of a Tangent Line

If the point $P(x_0,y_0)$ is on the curve f, then the tangent line at P has a slope that is given by

$$m_{tan} = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$$

provided that the limit exist.

Recall that the equation of the tangent line through point (x_0, y_0) with slope m is the point-slope form of a line: $y - y_0 = m_{\tan}(x - x_0)$.

Example 1: Find line tangent to the curve $f(x)=x^3$ that passes through point P (2,8).

Solution: Since $P(x_0, y_0) = (2, 8)$, using the slope of the tangent equation we have

Table 7.28:

	m_{tan}	$= \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$
and we get		
	m_{tan}	$= \lim_{h \to 0} \frac{f(2+h)-f(2)}{h}$ $= \lim_{h \to 0} \frac{(h^3+6h^2+12h+8)-8}{h}$ $= \lim_{h \to 0} \frac{h^3+6h^2+12h}{h}$ $= \lim_{h \to 0}(h^2 + 6h + 12)$ $= 12$

Thus the slope of the tangent line is 12. Using the point-slope formula above, we find that the equation of the tangent line is y -8 = 12 (x -2) or y = 12 x -16.

What we are interested in next is to be able to know the slope of the tangent line at *any point* on the curve f. Such a formula would be the same formula that we are using except we would replace the constant x_0 by the variable x. This yields,

Table 7.29:

	m_{tan}	$= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$
We denote this formula by,		
	$f'(x)$	$= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$

where $f'(x)$ reads "f prime of x" The next example illustrate its usefulness.

Example 2: If $f(x) = x^2 - 3$,find $f'(x)$ and use the result to find the slope of the tangent line at $x = 2$ and $x = -1$.

Solution: Since $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$then

Table 7.30:

	$f'(x)$	$= \lim_{h\to 0} \frac{\left[(x+h)^2-3\right]-\left[x^2-3\right]}{h}$
		$= \lim_{h\to 0} \frac{x^2+2xh+h^2-3-x^2+3}{h}$ $= \lim_{h\to 0} \frac{2xh+h^2}{h}$ $= \lim_{h\to 0}(2x+h)$ $= 2x$

To find the slope,we simply substitute $x = 2$ into the result $f'(x)$:

Table 7.31:

	$f'(x)$	$= 2x$
and	f'(2)	=2(2) = 4
	f'(x)	=2x
	f'(-1)	=2(-1) = -2

Thus slope of the tangent line at $x = 2$ and $x = -1$ are 4 and -2 respectively.

Example 3: Find the slope of the tangent line to the curve $y = 1/x$ that passes through the point (1, 1).

Solution: Using the slope of the tangent formula,

Table 7.32:

	$f'(x)$	$= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$
and substituting $y = \frac{1}{x}$		
	$f'(x)$	$= \lim_{h\to 0} \frac{\left(\frac{1}{x+h}\right)-\frac{1}{x}}{h}$
		$= \lim_{h\to 0} \frac{\frac{x-x-h}{x(x+h)}}{h}$
		$= \lim_{h\to 0} \frac{x-x-h}{hx(x+h)}$
		$= \lim_{h\to 0} \frac{-h}{hx(x+h)}$
		$= \lim_{h\to 0} \frac{-1}{x(x+h)}$
		$= \frac{-1}{x^2}$
For x=1, the slope is		
	$f'(x)$	$= \frac{-1}{1} = -1$
		$= -1$

Thus the slope of the tangent line at $x = 1$ for the curve $y = 1/x$ is $m = -1$. To find the equation of the tangent line, we simply use the point-slope formula,

Table 7.33:

	y-y_0	$=m(x$-$x_0)$
Where $(x_0, y_0) = (1, 1)$.		
	y- 1	= -1(x-1)
	y	= -x+1+1
	y	= -x+2

So the equation of the tangent line is y=-x+2.

Average Rate of Change and Instantaneous Rate of Change

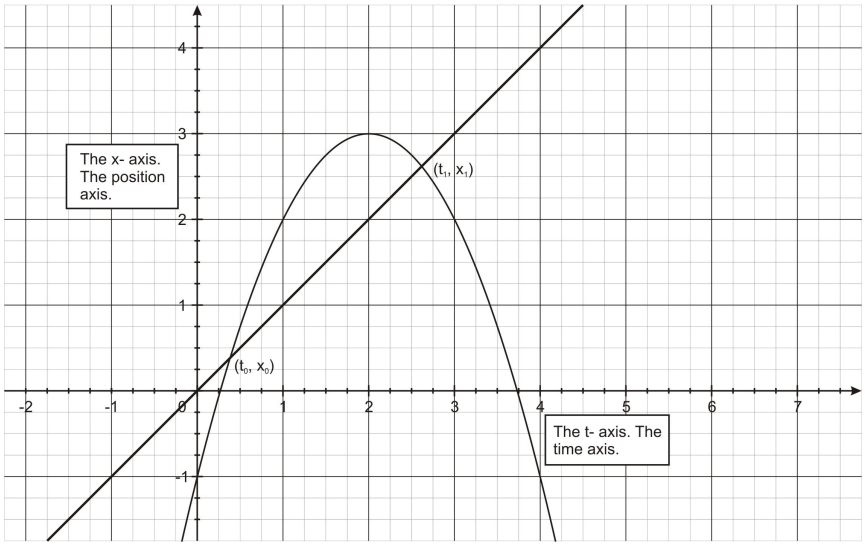
One of the two primary concepts of calculus involves calculating the rate of change of one quantity with respect to another. For example, speed is defined as the rate of displacement with respect to time. If a person travels 120 miles in 4 hours, his speed is 120/4=30 mi/hr. This speed is called the *average speed* or the *average rate of change* of distance with respect to time. Of course the person who travels 120 miles at a rate of 30 mi/hr for 4 hr probably does not do so continuously. Though he probably slowed down or sped up during the 4-hour period, it does suffice to say that he traveled for 4 hours at an average rate of 30 miles per hour. However, if the driver strikes a tree, it would not be his average speed that determines his survival but his speed at the *instant of the collision*. Similarly, when a bullet strikes a target, it is not the average speed that is significant but its *instantaneous speed* at the moment it strikes. So here we have distinct kinds of speeds, average speed and instantaneous speed.

The average speed of an object is defined as the object's displacement x divided by the time interval t during which the displacement occurs:

Table 7.34:

Average speed $= v = \frac{\Delta x}{\Delta t} = \frac{x_1-x_0}{t_1-t_0}$

Notice that the points(t_0, x_0) and(t_1, x_1) lie on the position versus time curve, as the figure below shows.



This expression is also the expression for the slope of a secant line connecting the two points. Thus we conclude that the average velocity of an object between time t_0 and t_1 is respresented geometrically by the slope of the secant line connecting the two points (t_o, x_o) and (t_1, x_1). If we choose t_1 close to t_o, then the average velocity will closely approximate the instantaneous velocity at time t_o.

Geometrically, the average rate of change is represented by the slope of a secant line (figure a, below) and the instantaneous rate of change is represented by the slope of the tangent line (figure b, below).

Table 7.35:

Average Rate of Change(such as *the average velocity*) The average rate of change of $y=f(x)$ over the time interval $[x_0, x_1]$ is the slope m_{sec} of the secant line to the points $(x_o, f(x_0))$ and $(x_1, f(x_0))$ on the graph (figure a):

$$m_{sec} = \frac{f(x_1)-f(x_0)}{x_1-x_0}$$

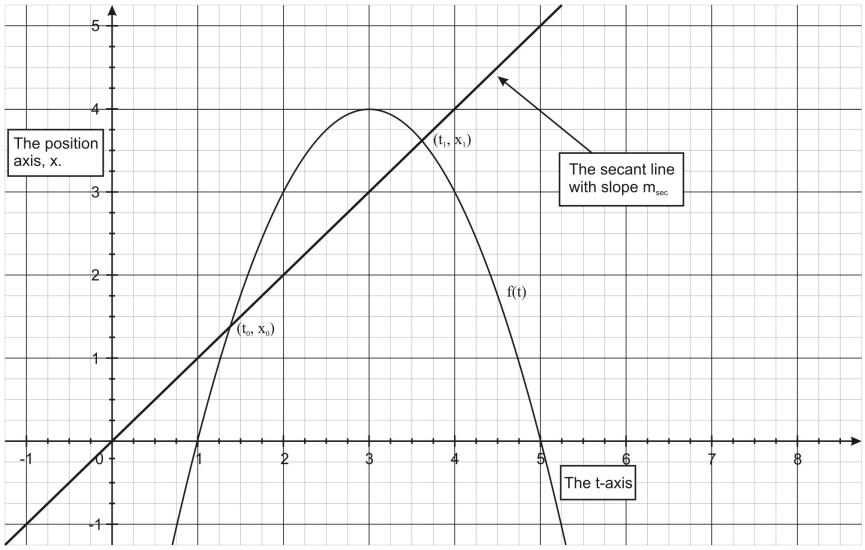
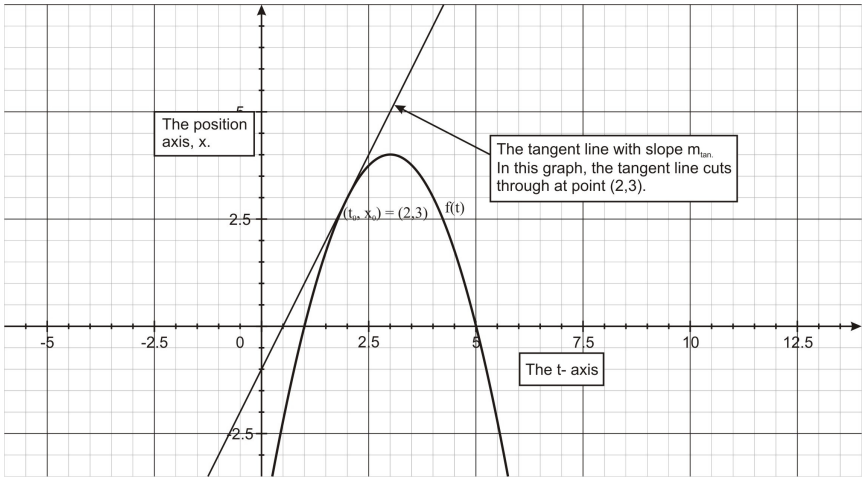


Table 7.36:

Instantaneous Rate of Change The instantaneous rate of change of $y=f(x)$ at the point x_0 is the slope m_{sec} of the tangent line to the point x_0 on the graph (figure b):

$$m_{tan} = f'(x_0) = \lim_{x_1 \to x_0} \frac{f(x_1)-f(x_0)}{t_1-t_0}$$



Example 4: Suppose that $y = x^2 - 3$.

1. Find the average rate of change of y with respect x over the interval [0, 2].
2. Find the instantaneous rate of change of y with respect x at the point $x = -1$.

Solution:

1. Applying the formula above for secant with $f(x)=x^2$-3 and $x_0=0$ and $x_1=2$, yields

Chapter 8

Sequences, Series, and Mathematical Induction

8.1 Recursive and Explicit Formulas

Learning objectives

- Write a recursive formula for a sequence, and use the formula to identify terms in the sequence.
- Write an explicit formula for a sequence, and use the formula to identify terms in the sequence.
- Identify a sequence as arithmetic, geometric, or neither.

Introduction

Consider a situation in which the value of a car depreciates 10% per year. If the car is originally valued at $20,000, the following year it is worth 90% of $20,000, or $18,000. After another, the value is 90% of $18,000, or $16,200. If we write the decreasing values as a list: 20,000, 18,000, 16,200 ... we have written a **sequence.**

A sequence is an ordered list of objects. Above our "objects" were all numbers. The simplest way to represent a sequence is by listing some of its terms. For example the sequence of odd, positive integers is shown here:

Table 8.1:

1, 3, 5, 7 ...

In this lesson you will learn to represent a sequence in two ways. The first method requires that you know the previous term in order to find the next term in the sequence. The second method does not.

Representing a Sequence Recursively

Consider again the sequence shown above. What is the next term?

As long as you are familiar with the odd integers (i.e., you can count in 2's) you can figure out that the

next term is 9. If we want to describe this sequence in general, we can do so by stating what the first term is, and then by stating the relationship between successive terms. When we represent a sequence by describing the relationship between its successive terms, we are representing the sequence **recursively**.

The terms in a sequence are often denoted with a variable and a subscript. All of the terms in a given sequence are written with the same variable, and increasing subscripts. So we might list terms in a sequence as a_1, a_2, a_3, a_4, a_5 ...

We can use this notation to represent the example above. This sequence is defined as follows:

Table 8.2:

$a_1 = 1$
$a_n = a_{n\text{-}1} + 2$

At first glance this notation may seem confusing. What is important to keep in mind is that the subscript of a term represents its "place in line." So a_n just means the n^{th} term in the sequence. The term $a_{n\text{-}1}$ just means the term before a_n. In the sequence of odd numbers above, $a_1 = 1$, $a_2 = 3$, $a_3 = 5$, $a_4 = 7$, $a_5 = 9$ and so on. If, for example, we wanted to find a_{10}, we would need to find the 9^{th} term in the sequence first. To find the 9^{th} term we need to find the 8^{th} term, and so on, back to a term that we know.

Example 1: For the sequence of odd numbers, list a_6, a_7, a_8, a_9, and a_{10}

Solution: Each term is two more than the previous term.

$a_6 = a_5 + 2 = 9 + 2 = 11$
$a_7 = a_6 + 2 = 11 + 2 = 13$
$a_8 = a_7 + 2 = 13 + 2 = 15$
$a_9 = a_8 + 2 = 15 + 2 = 17$
$a_{10} = a_9 + 2 = 7 + 2 = 19$

Finding terms in this sequence is relatively straightforward, as the pattern is familiar. However, this would clearly be tedious if you needed to find the 100^{th} term. We will turn to another method of defining sequences shortly. First let's consider some more complicated sequences.

Example 2: For each sequence, find the indicated term.

a. Find the 5^{th} term for the sequence:

Table 8.3:

$t_1 = 3$
$t_n = 2t_{n\text{-}1}$

b. Find the 4^{th} term for the sequence:

Table 8.4:

$b_1 = 3$
$b_n = (b_{n\text{-}1})^2 + 1$

Solution:

a. $t_5 = 48$

Table 8.5:

		$t_2 = 2t_1 = 2 \times 3 = 6$
$t_1 = 3$ $t_n = 2 \times t_{n\text{-}1}$	→	$t_3 = 2t_2 = 2 \times 6 = 12$ $t_4 = 2t_3 = 2 \times 12 = 24$ $t_5 = 2t_4 = 2 \times 24 = 48$

b. $b_4 = 677$

Table 8.6:

		$b_2 = (b_1)^2 + 1 = 2^2 + 1 = 4 + 1 = 5$
$b_1 = 2$ $b_n = (b_{n\text{-}1})^2 + 1$	→	$b_3 = (b_2)^2 + 1 = 5^2 + 1 = 25 + 1 = 26$ $b_4 = (b_3)^2 + 1 = 26^2 + 1 = 676 + 1 = 677$

As you can see from just a few terms of the sequence in example 2b, the terms in a sequence can grow quickly. If we compare the growth of the terms in the sequences we have seen so far, the first example, the sequence of odd numbers, was the slowest. Its growth is linear, and it is referred to as an **arithmetic sequence**. Every arithmetic sequence has a **common difference**, or a constant difference between each term. (The common difference is analogous to the slope of a line.) The sequence of odd numbers has a common difference of 2 because for all n, a_n - $a_{n-1} = 2$. The sequence in example 2a is a **geometric sequence**. Every geometric sequence has a **common ratio**. In the sequence in example 2a, the common ratio is 2 because for all n, $\frac{t_n}{t_{n-1}} = 2$. The terms of a geometric sequence follow an exponential pattern. The sequence in example 2b is neither arithmetic nor geometric, though its values follow a cubic pattern.

For any of these sequences, as noted above, determining more than a few values by hand would be time consuming. Next you will learn to define a sequence in a way that makes finding the n^{th} term faster.

Representing a sequence explicitly

When we represent a sequence with a formula that lets us find any term in the sequence without knowing any other terms, we are representing the sequence explicitly.

Given a recursive definition of an arithmetic or geometric sequence, you can always find an explicit formula, or an equation to represent the n^{th} term of the sequence. Consider for example the sequence of odd numbers we started with: 1,3,5,7,...

We can find an explicit formula for the n^{th} term of the sequence if we analyze a few terms:

$a_1 = 1$
$a_2 = a_1 + 2 = 1 + 2 = 3$
$a_3 = a_2 + 2 = 1 + 2 + 2 = 5$
$a_4 = a_3 + 2 = 1 + 2 + 2 + 2 = 7$

$a_5 = a_4 + 2 = 1 + 2 + 2 + 2 + 2 = 9$
$a_6 = a_6 + 2 = 1 + 2 + 2 + 2 + 2 + 2 = 11$

Note that every term is made up of a 1, and a set of 2's. How many 2's are in each term?

Table 8.7:

a_1	$= 1$
a_2	$= 1 + 2 = 3$
a_3	$= 1 + 2 \times 2 = 5$
a_4	$= 1 + 3 \times 2 = 7$
a_5	$= 1 + 4 \times 2 = 9$
a_6	$= 1 + 5 \times 2 = 11$

The n^{th} term has $(n$ - $1)2$'s. For example, $a_{99} = 1 + 98 \times 2 = 197$. We can therefore represent the sequence as $a_n = 1 + 2(n$ - $1)$. We can simplify this expression:

Table 8.8:

a_n	$=1+2(n\text{-}1)$
a_n	$=1 + 2n$ - 2
a_n	$=2n\text{-}1$

In general, we can represent an arithmetic sequence in this way, as long as we know the first term and the common difference, d. Notice that in the previous example, the first term was 1, and the common difference, d, was 2. The n^{th} term is therefore the first term, plus $d(n\text{-}1)$:

Table 8.9:

a_n	$=a_1+d(n\text{-}1)$

You can use this general equation to find an explicit formula for any term in an arithmetic sequence.

Example 3: Find an explicit formula for the nth term of the sequence 3,7,11,15... and use the equation to find the 50^{th} term in the sequence.

Solution: $a_n=4n\text{-}1$, and $a_{50}=199$

The first term of the sequence is 3, and the common difference is 4.

Table 8.10:

a_n	$=a_1+d(n\text{-}1)$
a_n	$=3+4(n\text{-}1)$
a_n	$=3+4n\text{-}4$
a_n	$=4n\text{-}1$
a_{50}	$=4(50)\text{-}1=200\text{-}1=199$

We can also find an explicit formula for a geometric sequence. Consider again the sequence in example 2a:

Table 8.11:

		$t_2 = 2t_1 = 2 \times 3 = 6$
$t_1 = 3$ $t_n = 2 \times t_{n\text{-}1}$	$\rightarrow$	$t_3 = 2t_2 = 2 \times 6 = 12$ $t_4 = 2t_3 = 2 \times 12 = 24$ $t_5 = 2t_4 = 2 \times 24 = 48$

Notice that every term is the first term, multiplied by a power of 2. This is because 2 is the common ratio for the sequence.

Table 8.12:

	t_1	$=3$
	t_2	$=2 \times 3 = 6$
	t_3	$=2 \times 2 \times 6 = 2^2 \times 6 = 12$
	t_4	$=2 \times 2 \times 2 \times 6 = 2^3 \times 6 = 24$
	t_5	$=2 \times 2 \times 2 \times 26 = 2^4 \times 6 = 48$

The power of 2 in the n^{th} term is (n-1). Therefore the n^{th} term in this sequence can be defined as: $t_n=3(2^{n\text{-}1})$. In general, we can define the n^{th} term of a geometric sequence in terms of its first term and its common ratio, r:

Table 8.13:

	t_{n}	$=t_1(\text{r}^{n\text{-}1})$

You can use this general equation to find an explicit formula for any term in a geometric sequence.

Example 4: Find an explicit formula for the n^{th} term of the sequence 5,15,45,135... and use the equation to find the 10^{th} term in the sequence.

Solution: $a_n=5 \times 3_1$,and $a_{10} = 98{,}415$

The first term in the sequence is 5, and $r = 3$.

Table 8.14:

	a_n	$=a_1 \times r^{n\text{-}1}$
	a_n	$=5 \times 3^{\text{n-1}}$
	a_{10}	$=5 \times 3^{10\text{-}1}$
	a_{10}	$=5 \times 3^9 = 5 \times 19{,}683 = 98{,}415$

Again, it is always possible to write an explicit formula for terms of an arithmetic or geometric sequence. However, you can also write an explicit formula for other sequences, as long as you can identify a pattern. To do this, you must remember that a sequence is a function, which means there is a relationship between the input and the output. That is, you must identify a pattern between the term and its **index**, or the term's "place" in the sequence.

Example 5: Write an explicit formula for the nth term of the sequence 1,(1/2),(1/3),(1/4)...

Solution: $a_n=(1/n)$

Initially you may see a pattern in the fractions, but you may also wonder about the first term. If you write 1 as (1/1), then it should become clear that the n^{th} term is $(1/n)$.

Lesson Summary

In this lesson you have learned to represent a sequence in three ways. First, you can represent a sequence just by listing its terms. Second, you can represent a sequence with a recursive formula. Given a list of terms, writing a formula requires stating the first term and the relationship between successive terms. Finally, you can write a sequence using an explicit formula. Doing this requires identifying a pattern between n and the n^{th} term.

In this lesson we have looked at several kinds of sequences: arithmetic, geometric, and sequences that do not follow either pattern. In the remainder of the chapter you will again see these three general categories of sequences.

Points to Consider

1. What is the difference between a recursive and an explicit representation of a sequence?
2. How many terms do sequences have?
3. What happens if we add up the terms in a sequence?

Review Questions

1. Find the value of a_6 , given the sequence defined as:

Table 8.15:

a_1	$=4$
a_n	$=5a_{n\text{-}1}$

2.
3. Find the value of a_5, given the sequence defined as:

Table 8.16:

a_1	$=32$
a_n	$=(1/2)a_{n\text{-}1}$

4.
5. Find the value of $a_{n\text{-}1}$, given the sequence defined as:

Table 8.17:

	a_1	$=1$
	a_n	$=3a_{n\text{-}1}\text{-}n$

6.
7. Consider the sequence: 2,9,16... Write an explicit formula for the sequence, and use the formula to find the value of the 20^{th} term.
8. Consider the sequence: 5,10,20... Write an explicit formula for the sequence, and use the formula to find the value of the 9_{th} term.
9. Consider the sequence (1/2)(1/4)(1/8) Write an explicit formula for the sequence, and use the formula to find the value of the 7_{th} term.
10. Identify all sequences in the previous six problems that are geometric. What is the common ratio in each sequence?
11. Consider the situation in the introduction: a car that is originally valued at $20,000 depreciates by 10% per year. What kind of sequence is this? What is the value of the car after 10 years?
12. In a particular arithmetic sequence, the second term is 4 and the fifth term is 13. Write an explicit formula for this sequence.
13. The membership of an online dating service increases at an average rate of 8% per year. In the first year, there are 500 members. a. How many members are there in the second year?
    b. How many members are there in the eighteenth year?

Review Answers

1. The 6^{th} term is 12,500
2. The 5^{th} term is 2
3. The 7^{th} term is -178
4.

Table 8.18:

	a_{n}	$=7n\text{-}5$
	a_{20}	$=135$

5.
6.

Table 8.19:

	a_{n}	$=5 \times 2^{n\text{-}1}$
	a_9	$=1280$

7.
8.

Table 8.20:

a_n	$= \frac{1}{2^n}$
a_7	$= \frac{1}{128}$

9.
10. The sequence in question 1 has $r = 5$.The sequence is question 2 has $r = 1/2$.
    The sequence in question 5 has $r = 2$.
    The sequence in question 6 has $r = 1/2$.
11. The sequence is a geometric sequence. The value of the car after 10 years is approximately $7748.
12. a_n=3n-2
13. a. 540 members b. Approximately 1,998 members

Vocabulary

Explicit formula An explicit formula for a sequence allows you to find the value of any term in the sequence.

Natural numbers The natural numbers are a subset of the integers: {1,2,3,4,5....}

Recursive formula A recursive formula for a sequence allows you to find the value of the n^{th} term in the sequence if you know the value of the $(n\text{-}1)^{\text{th}}$ term in the sequence.

Sequence A sequence is an ordered list of objects.

8.2 Summation Notation

Learning objectives

- Write a finite sum using sigma notation.
- Read sigma notation and evaluate a sum.
- Write an infinite sum using sigma notation.
- Find terms of a sequence and find sums using a graphing calculator.

Introduction

Consider the following situation:

You are participating in a 100-mile walk for a charity. In order to complete the entire walk, you first have to go halfway, or 50 miles. After you complete the first half of the walk, you again have to go halfway from where you are, or 25 more miles. Continuing with this pattern, you will always be going halfway. So will you ever finish the 100-mile walk?

This situation may seem silly, but it is actually a version of a problem that was first proposed thousands of years ago, by Zeno of Elea, a Greek philosopher who wanted to challenge accepted ideas about space and time. In today's world, most adults have an intuitive understanding of distance, rate and time, that allows us to answer the above question: yes, you will finish the walk, though it depends on how *fast* you walk.

We can also answer this question if we add up the sequence of distances you would walk each time you went half of the previous distance: 50+25+12.5+... even though this is the sum of infinitely many numbers, we can actually find the value of this expression. By the end of the chapter, you will be able to do this type of calculation.

An infinite sum is known as an infinite series. In this lesson you will learn about series, both finite and infinite. You will learn to use a particular kind of notation that allows you to write a sum succinctly.

Summation Notation

Consider for example a sequence defined by $a^n=3n$. *If we write out the sum of the first 4 terms, we have 3+6+9+12=30 . But what if we want to write out terms for a larger sum?*

Summation notation is a method of writing sums in a succinct form. To write the sum 3+6+9+12=30 , we use the Greek letter Sigma, as follows:

Table 8.21:

$\sum_{n=1}^{4} 3n$

The expression $3n$ is called the **summand**, the 1 and the 4 are referred to as the **limits** of the summation, and the n is called the **index** of the sum. Here we have used a "sigma" to write a sum. We can also read a sigma, and determine the sum. For example, we can read the above sigma notation as "find the sum of the first four terms of the series, where the n^{th} term is $3n$." We always read the limits from the bottom to the top. The bottom number tells you which term to start with, and the top limit tells you which term is the final term to add. We could then write out the sigma above as:

Table 8.22:

$\sum_{n=1}^{4} 3n$	=3(1)+3(2)+3(3)+3(4)
	=3+6+9+12=30

In general, we can either rewrite a given series in sigma notation, or we can read sigma notation in order to find the value of the sum.

Example 1: Write the sum using sigma notation:

2+4+6+8+10+12+14+16+19+20

Solution: $\sum_{n=1}^{10} 2n$

Every term is a multiple of 2. The first term is 2×1, the second term is 2×2 , and so on. So the summand of the sigma is $2n$. There are 10 terms in the sum. Therefore the limits of the sum are 1 and 10.

Properties of Sigma

Notice that we can write the sum 2 + 4 + 6 + 8 + 10 + 12 + 14 + 16 + 19 + 20 as 2(1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10) Therefore $\sum_{n=1}^{10} 2n = 2\sum_{n=1}^{10} n$. In general, we can factor a coefficient out of a sum:

Table 8.23:

$\sum_{n=1}^{k} ca_n$	$= \sum_{n=1}^{k} a_n$

We can also break a sum into two or more sums.

Example 2: Write out the terms of $\sum_{n=2}^{5}(n+7)$ and evaluate the sum.

Solution: The sum is 37.

Table 8.24:

$\sum_{n=2}^{5}(n+7)$	$=(2+7)+(3+7)+(4+7)+(5+7)$
	$=2+3+4+5+7+7+7+7$ $=2+3+4+5+7\times 4=14+28=42$

Notice we could have written $\sum_{n=2}^{5}(n+7)$ as $\sum_{n=2}^{5} n + \sum_{n=2}^{5} 7$. Also, the second sum does not depend on the index of the sum (i.e. it stays at 7 regardless of the index), only that there are 4 terms to add together. Seeing this can make a sum easier to evaluate.

In general, we can write a sum as a sum of sums: $\sum_{i-1}^{n}(a_i+b_i) = \sum_{i=1}^{n}(a_i) + \sum_{i=1}^{n}(b_i)$.

Example 3: Write out the terms of $\sum_{i=0}^{4} 32\left(\frac{1}{4}\right)^n$ and evaluate the sum.

Solution: The sum is $42\frac{5}{8}$

Table 8.25:

$\sum_{i=0}^{4} 32\left(\frac{1}{4}\right)^n$	$= 32\left(\frac{1}{4}\right)^0 + 32\left(\frac{1}{4}\right)^1 + 32\left(\frac{1}{4}\right)^2 + 32\left(\frac{1}{4}\right)^3 + 32\left(\frac{1}{4}\right)^4$
	$= 32\cdot 1 + 32\cdot\frac{1}{4} + 32\cdot\frac{1}{16} + 32\cdot\frac{1}{64} + 32\cdot\frac{1}{256}$ $= 32+8+2+\frac{1}{2}+\frac{1}{8} = 42\frac{5}{8}$

The properties of sigma we have used in the past 3 examples will be useful as we continue to work with sigma in this chapter, as well as in your future study of calculus. The same is true of the next topic, the infinite sum.

Infinite sums

Now let's return to the sum at the beginning of the lesson: 50 + 25 + 12.5 +...

We can write this sum using sigma notation if we identify the form of the n^{th} term. Notice that each term is half of the previous term. We can use this pattern to find the n^{th} term:

Table 8.26:

$a_1 = 50$
$a_2 = a_1 \cdot \frac{1}{2}$ $a_3 = a_2 \cdot \frac{1}{2} = a_1 \cdot \frac{1}{2} \cdot \frac{1}{2} = a_1 \left(\frac{1}{2}\right)^2$ $a_4 = a_1 \left(\frac{1}{2}\right)^3$ $a_n = a_1 \left(\frac{1}{2}\right)^{n-1}$

The terms form a geometric sequence, so we refer to the sum as an infinite geometric series. We can write this series as: $\sum_{n=1}^{\infty} 50\left(\frac{1}{2}\right)^{n-1}$. And, while this sum is the sum of infinitely many numbers, you can actually find the value of this sum. Remember: it is possible to complete the 100-mile charity walk. Therefore the sum is 100.

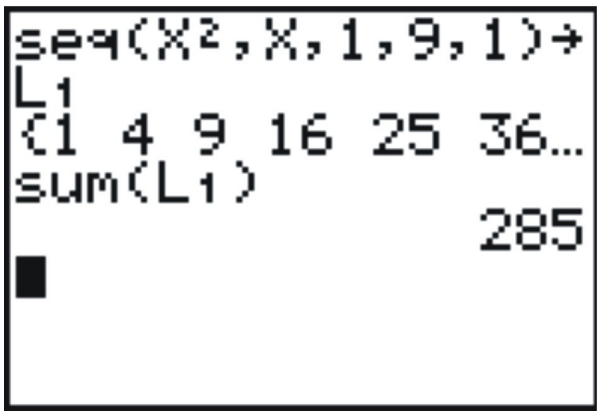
When an infinite sum has a finite value, we say the sum **converges.** Otherwise, the sum **diverges.** A sum converges only when the terms get closer to 0 after each step, but this is not a sufficient criterion for convergence. For example, the sum $\sum_{n-1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ...$ does not converge. Later in this chapter, you will learn how to find the value certain of infinite sums that converge.

Using a graphing calculator

To generate either a sequence or a series, you can use a graphing calculator. The TI-83/84 series gives you several options. You can, for example, work in sequence mode, which allows you to define a sequence and find terms. If you have an explicit formula for a sequence, you can keep your calculator in function mode. For example, consider the sum $\sum_{n=1}^{9} n^2$. It would be time consuming to write out the first 9 squares. The calculator is faster. To generate the 9 terms, press <TI font_2nd> [LIST], then select **OPS**, then option 5, **seq**(. This takes you back to the main screen. You should see **seq**(. After this, enter $x2$, x, 1, 9, 1). (The x tells the calculator that x is the input. The 1 and the 9 tell it the limits of the sum. The second 1 tells the calculator to go up in increments of 1.)

Press <TI font_ENTER>, and you should see the list of squares. Scroll to the right to see all of them. The scrolling will end when you reach 81.

If you want to find the sum of the terms, first store the sequence in a list (see screen below), then <TI font_2nd> [LIST], then select **MATH**, then option 5, **sum(.***Then* enter the name of the list and press <TI font_)><TI font_ENTER>. You should get 285.



Summary

In this lesson we have introduced Sigma notation, which allows us to write a sum without having to write out terms. We can use sigma notation to write a sum, and we can find the value of a sum, given its sigma notation. We can also represent an infinite sum using this notation. In some cases, an infinite sum has a finite value. Understanding the concepts in this lesson will help you to evaluate such sums later in the chapter.

Points to Consider

1. Consider the 100-mile walk example: how can infinitely many distances add up to 100?
2. What kind of infinite sum will converge? What will the terms look like?

Review Questions

For questions 1 – 3, express the sum using sigma notation

1. $1 + 3 + 5 + 7 + 9$
2. $2 + 6 + 18 + 54$
3. $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ... + \frac{1}{10}$
4. For questions 4 – 6, expand the sigma, and find the sum.
5. $\sum_{n=1}^{7}(2n-3)$
6. $\sum_{n=3}^{6}(n^2-5)$
7. $\sum_{n=1}^{5}\left(\frac{1}{3n}\right)$
8. Express the sum using sigma notation: $1 + 3 + 9 + 27 + ...$
9. Write the sum in question 4 as 2 sums.
10. Consider the sums $\sum_{n=1}^{5}(n+1)$ and $\sum_{n=1}^{5}(n-4)$. Is the product $\left(\sum_{n=1}^{5}(n+1)\right)\left(\sum_{n=1}^{5}(n-4)\right)$ equal to $\sum_{n=1}^{5}(n+1)(n-4)$?
11. Consider the sum $\sum_{n=1}^{\infty}\left(\frac{1}{5}\right)^n$. a. Find the sum of the first 10 terms using a graphing calculator.
    b. Find the sum of the first 20 terms using a graphing calculator
    c. Do you think the series converges? Explain.

Review Answers

1. $\sum_{n=1}^{5} 2n-1$
2. $\sum_{n=1}^{4} 2\left(3^{n-1}\right)$ or $\sum_{n=0}^{3} 2\left(3^n\right)$
3. $\sum_{n=1}^{10}\left(\frac{1}{n}\right)$
4.

Table 8.27:

$\sum_{n=1}^{7}(2n-3)$	$= (2\times1-3)+(2\times2-3)+(2\times3-3)+(2\times4-3)+(2\times5-3)+(2\times6-3)+(2\times7-3)$
	$= (-1)+(1)+(3)+(5)+(7)+(9)+(11)$ $= 35$

5.

6.

Table 8.28:

$\sum_{n=3}^{6}(n^2-5)$	$=(3^2-5)+(4^2-5)+(5^2-5)+(6^2-5)$
	$=(4)+(11)+(20)+(31)$
	$=66$

7.
8.

Table 8.29:

$\sum_{n=1}^{5}\left(\frac{1}{3n}\right)$	$=\left(\frac{1}{3}\right)+\left(\frac{1}{6}\right)+\left(\frac{1}{9}\right)+\left(\frac{1}{12}\right)+\left(\frac{1}{15}\right)$
	$=\frac{60}{180}+\frac{30}{180}+\frac{20}{180}+\frac{15}{180}+\frac{12}{180}$
	$=\frac{137}{180}$

9.
10. $\sum_{n=1}^{\infty}3^{n-1}$ or $\sum_{n=0}^{\infty}3^{n}$
11. $\sum_{n=1}^{7}(2n-3)=\sum_{n=1}^{7}(2n)+\sum_{n=1}^{7}(-3)$ or $2\sum_{n=1}^{7}(n)-\sum_{n=1}^{7}(3)$
12. The product of the sums is not equal to the sum of the product. $\sum_{n=1}^{5}(n+1)=20$ and $\sum_{n=1}^{5}(n-4)=-5$, so the product is -100
    $\sum_{n=1}^{5}(n+1)(n-4)=\sum_{n=1}^{5}\left(n^2-3n-4\right)=-10$
13. a. 0.2499999744 b. 0.25
    c. It likely converges to 0.25. If you find sums beyond the 20^{th} sum, the sum is still .25.

Vocabulary

Converge If the limit of the partial sums of series exists and is finite, the series converges.

Diverge If the limit of the partial sums of series does not exist or is infinite, then the series diverges.

Index The index of the sum is the variable in the sum.

Limits of the summation The limits of a summation are the starting and ending points of the sum.

Sigma Sigma (Σ) is the Greek letter used to represent a sum.

Summand The summand of a sigma is the expression being summed.

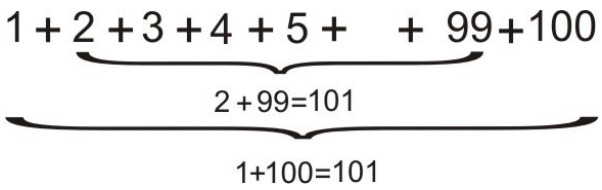
8.3 Mathematical Induction

Learning objectives

- Understand the process of mathematical induction.
- Use induction to prove integer sum formulas.

Introduction

In the previous lesson, you found sums of series with different numbers of terms. You probably noticed that adding together many numbers can be tedious, unless you use a calculator. Before people had computers and calculators, they often searched for ways to make calculations easier. For example, logarithms, which you learned about logarithms, were developed to simplify calculations that involve large numbers. Over time, mathematicians have developed formulas that allow us to simplify the calculations of large sums. In fact, German mathematician C.F. Gauss is often credited with discovering such a formula when he was a young child. The story is likely apocryphal (a legend), but it has been passed down since Gauss lived in the 1700's. According to the story, Gauss's teacher wanted to occupy students by having them add up large sets of numbers. When Gauss was asked to add up the first 100 integers, he found the sum very quickly, by pairing the numbers:



All of the numbers in the sum could be paired to make groups of 101. There are one hundred numbers being added, so there are such fifty pairs. Therefore the sum is $50(101) = 5050$.

The method Gauss used to solve this problem is the basis for a formula that allows us to add together the first n positive integers. However, in order to prove that the formula always works, we need to show that it works for all positive integers. In this lesson you will learn about **mathematical induction**, a method of proof that will allow you to prove that a particular statement is true for all positive integers. First we will present the method, and then we will prove Gauss's formula, as well another related sum.

The Method of Induction

First let's make a guess at a formula that will give us the sum of all the positive integers from 1 to n for any integer n. If we look closely at Gauss's method, we can see a general form: there were 100 numbers, hence 50 pairs. So if there were n numbers, there would be $(n/2)$ pairs. The first and last numbers were 1 and 100. They added together to give us 101. This number was the sum of each pair in the overall sum. So in general, we could add together 1 and n to get the sum of each pair. Therefore we might hypothesize that the sum of the first n positive integers is $n((1 + \mathrm{n})/2)$. However, we have not proven that this formula works for all positive integers n. Mathematical induction will allow us to do this. The overall idea of induction is this: you assume that a statement is true for some arbitrary value of n (often it is called k), and then you show that if the statement is true for $n{=}k$, it must also be true for $n{=}k{+}1$. We do this precisely because we can't actually show it is true for every value. For example, you might show that the above equation is true for $n = 100$, and then $n = 101$, and then $n = 102$, but then what about 103? 104? 500? A million?

Mathematical induction allows us prove that a statement is true in three steps:

1. *The base case:* prove that the statement is true for the first value of n. In some cases, this might be $n = 0$. In the case of the integer sum formula above, we would start with $n = 1$. Often with induction you may want to expand the first step by showing that the statement is true for several values of n.
2. *The inductive hypothesis:* assume that the statement is true for the k^{th} value of n. In the case of the integer sum formula, we would state the following: the sum of the first k positive integers is $k((1 + k)/2)$.

3. *The inductive step:* use the inductive hypothesis to show that the statement is true for the $k + 1^{\text{th}}$ step. In the case of the integer sum formula, we would prove the following: assuming that the sum of the first k positive integers is $k((1 + k)/2)$, the sum of the first k +1 positive integers is $((k + 1)$ $(1 + (k + 1)/2)$.

Carrying out this kind of proof requires that you perform each of these steps. In particular, for the third step you must rely on your algebra skills. Next we will prove Gauss's formula as an example of carrying out induction.

Proof of the sum of the first n integers

Prove: The sum of the first n positive integers is $\frac{n(1+n)}{2}$.

1. The base case:

If $n = 1$, the entire sequence is just 1 and therefore the sum is 1. Also, $\frac{n(1+n)}{2} = \frac{1(1+1)}{2} = \frac{2}{2} = 1$. This establishes the base case.

2. Assume that the sum of the first k positive integers is $\frac{k(1+k)}{2}$. In other words, assume that $1 + 2 + 3 + \cdots + k = \frac{k(1+k)}{2}$.

3. We must show that the sum of the first $k+1$ positive integers is $\frac{(k+1)(1+(k+1))}{2}$. In other words, we must show that

$1 + 2 + 3 + \cdots + k + (k + 1) = \frac{(k+1)(1+(k+1))}{2}$.

There are two key ideas to keep in mind as you are carrying out this step: (1) remember to use the assumption and (2) remember how sums work.

How does the sum of the first $k + 1$ integers relate to the sum of the first k integers? To get the sum of the first $k + 1$ integers we must add up all the integers from 1 to k and then add on $k + 1$, since the sum of the first $k + 1$ integers is $1 + 2 + 3 + \cdots + k + (k + 1)$.

Now we must use our assumption. Remember that we are *assuming* that $1 + 2 + 3 + \cdots + k = \frac{k(1+k)}{2}$. So we can substitute $\frac{k(1+k)}{2}$ in for $1 + 2 + 3 + \cdots + k$ in our expression above for the sum of the first $k +$ 1 integers. Now we have the sum of the first $k + 1$ integers is $\frac{k(1+k)}{2} + (k + 1)$.

Remember that we are trying to show that the sum of the first $k + 1$ integers is $\frac{(k+1)(1+(k+1))}{2}$. With some algebraic manipulation, we can show that $\frac{k(1+k)}{2} + (k + 1) = \frac{(k+1)(1+(k+1))}{2}$. See below:

Table 8.30:

$\frac{k(1+k)}{2} + (k + 1)$	
$= \frac{k(k+1)}{2} + \frac{2(k+1)}{2}$	The common denominator is 2 Add the fractions
$= \frac{k(k+1)+2(k+1)}{2}$	Simplify the numerator
$= \frac{k^2+k+2k+2}{2}$	

Table 8.30: (continued)

$\frac{k(1+k)}{2} + (k+1)$	
$= \frac{k^2+3k+2}{2}$	Factor the numerator
$= \frac{(k+1)(k+2)}{2}$	The term $(k+2)$ is the same as $((k+1)+1)$
$= \frac{(k+1)((k+1)+1)}{2}$	

We have shown that our formula for the sum of the first n integers is true for n=1. We have also shown that whenever it is true for n=k it is also true for n=k+1. Since we know it is true for n=1, it must therefore be true for n=2. Similarly, since it is true for n=2, it must therefore be true for n=3, and it must therefore be true for n=4,... You should see that we have proven that the sum of the first n positive integers is $\frac{n(1+n)}{2}$ for *all* integer values of n. We can similarly prove a formula for the sum of the first n terms in an arithmetic series.

The nth partial sum of an arithmetic series

We can use induction to prove that the sum of the first n terms of an arithmetic series is $S_n = \frac{n(a_1+a_n)}{2}$, where a_1 is the first term in the series and a_n is the last term.

Recall that in an arithmetic sequence or series, there is a common difference, d, between each term, and that the n$^{\text{th}}$ term is $a_n = a_1 + d(n-1)$. We need to keep these ideas in mind in order to complete the proof.

1. Base case: if $n = 1$, then $S_1 = a_1$. Using the hypothesized formula, we have

Table 8.31:

$S_1 = \frac{1(a_1+a_1)}{2} = \frac{2a_1}{2} = a_1$

2. Assume that $S_k = \frac{k(a_1+a_k)}{2}$.

3. Prove that if our formula for S_k is true then $S_{k+1} = \frac{(k+1)(a_1+a_{k+1})}{2}$.

We can think of the sum of the first $k+1$ terms as the sum of the first k terms, plus the $k+1$ term. So we have:

Table 8.32:

$S_{k+1} = S_k + a_{k+1}$	Add the $k+1$ term
$= \frac{k(a_1+a_k)}{2} + a_{k+1}$	Use the formula for S_k from step 2
$= \frac{k(a_1+a_k)}{2} + \frac{2a_{k+1}}{2}$	The common denominator is 2
$= \frac{k(a_1+a_k)+2a_{k+1}}{2}$	Add the fractions
$= \frac{k(a_1+(a_1+(k-1)d))+2(a_1+kd)}{2}$	Use substitution: remember that $a_m = a_1+(m-1)d$ for any positive integer m. So $a_k = a_1 + (k-1)d$ and $a_{k+1} = a_1 + (k)d$
$= \frac{k(a_1+a_1+kd-d)+2a_1+2kd}{2}$	
$= \frac{ka_1+ka_1+k^2d-kd+2a_1+2kd}{2}$	Distribute and combine like terms
$= \frac{2ka_1+2a_1+k^2d+kd}{2}$	Factor by grouping

Table 8.32: (continued)

$S_{k+1} = S_k + a_{k+1}$	Add the $k+1$ term
$= \frac{2a_1(k+1)+kd(k+1)}{2}$ $= \frac{(k+1)(2a_1+kd)}{2}$ $= \frac{(k+1)(a_1+a_1+kd)}{2}$ $= \frac{(k+1)(a_1+a_{k+1})}{2}$	Again, $a_{k+1} = a_1 + (k)d$

So we have now proven that the sum of the first n terms in an arithmetic series is $S_n = \frac{n(a_1+a_n)}{2}$. We can use this equation as a formula to find a sum.

Example: Find the sum of the first 50 terms of an arithmetic series if the first term is 5 and the common difference is 3.

Solution:

$S_{50} = \frac{50(a_1+a_n)}{2} = \frac{50(5+(5+49\times3))}{2} = \frac{50(157)}{2} = (25)(127) = 3,175$

This method is clearly much easier than writing out and adding 50 numbers!

Summary

In this lesson we have introduced the technique of mathematical induction. This method of proof allows you to prove that a statement is true for all positive integers. In this lesson we have focused on statements involving sums: we proved a formula for the sum of the first n positive integers, and a formula for the sum of the first n terms in an arithmetic series. In the next lesson we will use induction to prove other kinds of statements involving integers.

Points to Consider

1. What is k?
2. How is induction different from other forms of proof?
3. What else can we prove with induction?

Review Questions

1. Use Gauss's formula to find the sum of the first 200 positive integers.
2. If the sum of the first n integers is 210, what is n?
3. Find the sum of the first *40* terms of an arithmetic series in which the first term is *8* and the common difference is 5.
4. The sum of the first 28 terms in an arithmetic series is 1,946. If the first term is 2, what is the hundredth term?
5. Consider the series: -13 + -3 + 7 + 17 + 27 + ... a. What is the 25^{th} term?
   b. What is the sum of the first 25 terms?
6. A student needed to find the sum of the first 10 terms of the series 4 + 12 + 36 + ... and so he wrote the following: $S_{10} = \frac{10(4+4\times3^9)}{2} = \frac{10(78,736)}{2} = 393,680$
   Do you agree with the student's work? Explain.
7. Use induction to prove that $1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$

8. Use induction to prove that $1+3+5+\cdots+(2n-1)=n^2$
9. Use induction to prove that $1^3+2^3+3^3+\cdots+n^3=\frac{n^2(n+1)^2}{4}$
10. Use induction to prove that $1+4+7+\cdots+(3n-2)=\frac{3n^2-n}{2}$.

Review Answers

1. 20,100
2. $n=20$
3. 4,220
4. 137
5. a. 227 b. 2,675
6. Because the series is geometric, this formula is not appropriate. The work here does not represent the sum of the first 10 terms. Using a graphing calculator, you can find that the sum is 118,096.
7. 1. Base case: $1^2=1$
   $\frac{1(1+1)(2(1)+1)}{6}=\frac{2(3)}{6}=1$
   2. Inductive hypothesis: $1^2+2^2+3^2+...+k^2=\frac{k(k+1)(2k+1)}{6}$
   3. Inductive step: show that $1^2+2^2+3^2+...+k^2+(k+1)^2=\frac{(k+1)(k+1+1)(2(k+1)+1)}{6}$
   First, note that $\frac{(k+1)(k+1+1)(2(k+1)+1)}{6}=\frac{(k+1)(k+2)(2k+3)}{6}$.
   Now we have:

Table 8.33:

	$1^2+2^2+3^2+...+k^2+(k+1)^2$
	$=\frac{k(k+1)(2k+1)}{6}+(k+1)^2$ $=\frac{k(k+1)(2k+1)+6(k+1)^2}{6}=\frac{(k+1)[k(2k+1)+6(k+1)]}{6}$ $=\frac{(k+1)\left[2k^2+k+6k+6\right]}{6}=\frac{(k+1)\left[2k^2+7k+6\right]}{6}=\frac{(k+1)(2k+3)(k+2)}{6}$

8.
9. 1. Base case: $1=1^2$2. Inductive hypothesis: assume that $1+3+5+...+(2k-1)=k^2$
   3. Show that $1+3+5+...+(2k-1)+(2k+1)=(k+1)^2$
   We have:
   $1+3+5+...+(2k-1)+(2k+1)=k^2+(2k+1)$

Table 8.34:

	$=k^2+2k+1$
	$=(k+1)(k+1)=(k+1)^2$

10.
11. 1. Base case: $1^3=1$
    $\frac{1^2(1+1)^2}{4}=\frac{2^2}{4}=1$
    2. Assume that $1^3+2^3+3^3+...+k^3=\frac{k^2(k+1)^2}{4}$
    3. Show that $1^3+2^3+3^3+...+k^3+(k+1)^3=\frac{(k+1)^2((k+1)+1)^2}{4}$
    First, note that $\frac{(k+1)^2((k+1)+1)^2}{4}=\frac{(k+1)^2(k+2)^2}{4}$.
    Now we have:

Table 8.35:

	$1^3 + 2^3 + 3^3 + ... + k^3 + (k+1)^3$
	$= \frac{k^2(k+1)^2}{4} + (k+1)^3$ $= \frac{k^2(k+1)^2+4(k+1)^3}{4}$ $= \frac{(k+1)^2\left[k^2+4(k+1)\right]}{4} = \frac{(k+1)^2\left[k^2+4k+4\right]}{4} = \frac{(k+1)^2(x+2)^2}{4}$

12.

13. 1. Base case: $1 = 1\ \frac{3(1)^2-1}{2} = \frac{2}{2} = 1$
    2. Inductive hypothesis: assume that $1 + 4 + 7 + ... + (3k-2) = \frac{3k^2-k}{2}$
    3. Show that $1 + 4 + 7 + ... + (3(k+1)-2) = \frac{3(k+1)^2-(k+1)}{2}$

    First, note that

Table 8.36:

	$1 + 4 + 7 + ... + (3(k+1)-2) = \frac{3(k+1)^2-(k+1)}{2}$
	$= \frac{(k+1)[3(k+1)-1]}{2}$ $\frac{(k+1)[3k+2]}{2}$

14.

15. Now we have:

Table 8.37:

	$1 + 4 + 7 + ... + (3k-2) + (3(k+1)-2)$
	$= \frac{3k^2-k}{2} + (3(k+1)-2)$ $= \frac{3k^2-k}{2} + (3k+1)$ $\frac{3k^2-k+2(3k+1)}{2} = \frac{3k^2-k+6k+2}{2} = \frac{3k^2+5k+2}{2} = \frac{(3k+2)(k+1)}{2}$

16.

Vocabulary

Hypothesis A hypothesis is a conclusion made on the basis of evidence, or a statement assumed to be true for the sake of an argument.

Mathematical Induction Mathematical induction is a method of mathematical proof used to establish that a given statement is true of all positive integers (natural numbers).

Partial sum A partial sum is the sum of the first n terms in an infinite series, where n is some positive integer.

8.4 Mathematical Induction, Factors, and Inequalities

Learning objectives

- Use induction to prove statements involving factors.
- Use induction to prove statements involving inequalities.

Introduction

In the previous lesson we introduced mathematical induction, a method that allows us to prove that a given statement is true for all positive integers. The proofs in the previous lesson focused on sums. For example, we proved a formula for the sum of the first n positive integers. In this lesson we will use induction to prove other kinds of statements. First we will prove statements about factors, and then we will prove statements about inequalities.

First we discuss several important properties of integers that will be used to prove statements about factors.

Integers and Factors

There are two properties of integers and their factors that will be useful for the proofs in this lesson.

Property 1: If a is a factor of b, and a is a factor of c, then a is a factor of the sum $b+c$.

For example, the set of numbers $a = 3$, $b = 6$, and $c = 9$ satisfies the hypothesis because 3 is a factor of 6 and 3 is a factor of 9. Property 1 states that therefore 3 is a factor of $9 + 3 = 12$, which we know is true because $3 \times 4 = 12$. We can prove that this property is true for all integers if we think about what the term *factor* means. If a is a factor of b, then there exists some integer M such that $aM=b$. Similarly if a is a factor of c, then there exists some integer N such that $aN=c$. So we can write the sum $b+c$ as $aM+aN$. We know

Table 8.38:

$b + c = aM + aN = a(M + N)$.

Because we can write the sum as a product of a and another number, a is a factor of the sum $b + c$.

Property 2: If a is a factor of b and b is a factor of c, then a is a factor of c.

We can prove this property in a similar manner. If a is a factor of b, then there exists some integer M such that $aM=b$. If b is a factor of c, then there exists some integer N such that $bN=c$. We can write c in the following manner:

$c=bN=(aM)N=a(MN)$

Therefore a is a factor of c because we can write c as the product of a and another integer, MN.

As noted above, these properties of integers are useful for proving statements about integers and factors via induction.

Induction and factors

Here we will use induction to prove statements about factors. The first example focuses on integers. The second focuses on polynomials.

Example 1: Prove that 3 is a factor of 4^n-1 for all positive integers n.

Proof by induction:

1. The base case: if $n = 1$, then 4^n-1=4-1=3. 3 is a factor of itself because $3 \times 1=3$.
   If the base case does not convince you, you can always test out additional values of n. For example, if $n = 2$, 4^2-1=16-1=15=5×3.
2. The inductive hypothesis: assume that 3 is a factor of 4^k-1.
3. The inductive step: show that 3 is a factor of 4^{k+1}-1.
   If 3 is a factor of 4^k-1, then there exists some integer M such that $3M=4^k$-1. We can write 4^{k+1}-1 in a manner that allows us to use the inductive hypothesis:

Table 8.39:

	4^{k+1}-1	
	$4(4^k\text{-}1)+3$	Factor out 4, but add 3 in order to keep the value of 4^{k+1}-1
	$4(3M)+3$	Substitute: $3M=4^k$-1

   Note that the substitution is the same as using property 2 above: if 3 is a factor of 4^k-1 , then 3 is a factor of $4(4^k\text{-}1)$. Using the substitution simply makes the fact a bit more obvious.
   This last step proves that 3 is a factor of 4^{k+1}-1 , by property 1 above.

The technique of rewriting the k+1 term can also be used to prove statements about polynomials and factors.

Example 2: Prove that x-y is a factor of x^n-y^n for all positive integers n.

Note: Since we are talking about polynomials that are factorable now, not integers, then we say that if x-y is a factor of x^n-y^n, then there exists a polynomial P such that $P(x\text{-}y)=x^n\text{-}y^n$.

Proof by induction:

1. The base case: If $n = 1$, we have x^n-y^n=x - y , and x - y is a factor of itself, as x - y=1(x - y).
   As we did above, we can also check $n = 2$ in order to convince ourselves. If $n = 2$, we have x^2-y^2=(x - y)(x + y), so x - y is clearly a factor.
2. The inductive hypothesis: assume that x - y is a factor of x^k-y^k.
3. Show that x - y is a factor of x^{k+1}-y^{k+1}.
   From the inductive step, we know that there is some polynomial P such that $P(x$ - $y)=x^k$-y^k. We can rewrite x^{k+1}-y^{k+1} in a manner that allows use to use the inductive hypothesis:

Table 8.40:

	x^{k+1}-y^{k+1}
	$=x^{k+1}$-xy^k + xy^k -y^{k+1} $=x(x^k$-$y^k)$ + $y^k(x$-$y)$ $=x(P(x$ - $y))$ + $y^k(x$ - $y)$ $=Px(x$ - $y)+y^{k-1}$(x - y)

Again, by property 1 above, this shows that x - y is a factor of x^{k+1}-y^{k+1}. Therefore we have shown that x - y is a factor of x^n-y^n for all positive integers n.

The two proofs here have established properties of integers and polynomials involving factors. Now we will use induction to prove several statements about inequalities.

The transitive property of inequality

Below, we will prove several statements about inequalities that rely on the transitive property of inequality:

Table 8.41:

	If $a < b$ and $b < c$, then $a < c$.

Note that we could also make such a statement by turning around the relationships (i.e., using "greater than" statements) or by making inclusive statements, such as a b.

It is also important to note that this property of integers is a **postulate**, or a statement that we assume to be true. This means that unlike the factor properties we used above, we need not prove the transitive property of inequality.

You encountered other useful properties of inequalities in earlier algebra courses:

Addition property: if $a > b$, then $a + c > b + c$.

Multiplication property: if $a > b$, and $c > 0$ then $ac > bc$.

We will use these properties in the proof below.

Induction and Inequalities

Example 3: Prove that $n!$ 2^n for n 4

Proof by induction:

1. The base case is n = 4:4!=24, 2^4=16. 24 16 so the base case is true.
2. Assume that $k!$ 2^k for some value of k such that k 4
3. Show that $(k+1)!$ 2^{k+1}

Table 8.42:

	$(k+1)!=k!(k+1)$	Rewrite $(k+1)!$ in terms of k !
	$2^k(k+1)$	Use step 2 and the multiplication property.
	$2^k(2)$	$k+1 \quad 5 > 2$, so we can use the multiplication property again.
	$=2^{k+1}$	

Therefore $n! \quad 2^n$ for $n \quad 4$.

Lesson Summary

In this lesson we have extended the method of mathematical induction beyond proofs involving sums. Here we have proven statements about factors, and statements about inequalities. Induction is the method of choice for such proofs because we need to prove that a given statement is true for all positive integers. However, induction is only one method of proof. Many mathematical relationships can be proven without induction.

Points to Consider

1. How are the statements proven in this lesson different from the ones proven in the previous lesson?
2. Why do we limit the values of n to the natural numbers?
3. How can we prove relationships without induction?

Review Questions

1. Without adding, determine if 7 a factor of 49+70.
2. Consider the sum 23+54=77. Is 7 a factor of 77? What does this tell you about the first factor property in the lesson?
3. For what values of x is the inequality $x > x^2$ true?
4. Evaluate each expression: a.4!+3!
   b. $\frac{6!}{5!}$
5. Prove that 9^n-1 is divisible by 8 for all positive integers n.
6. Prove that x^n-1 is divisible by x-1 for all positive integers n.
7. Prove that n^2-n is even for all positive integers n.
8. Prove that $5^{2n-1}+1$ is divisible by 6 for all positive integers n.
9. Prove that $2n+1 < 2^n$ for all integers $n > 3$
10. Prove that $3^n > n^2$ for all positive integers n.

Review Answer

1. 7 is a factor of the sum, as it is a factor of 49 and a factor of 70.
2. 7 is a factor of 77, but it is not a factor of 23 or 54. This tells us that the converse of the property is not necessarily true.

3. The inequality is true if x is a number between -1 and 1 but not 0.
4. a. 30 b. 6
5.

Table 8.43:

1. Base case:	If $n = 1$,9^n-1=9-1=8=8(1)
2. Inductive hypothesis:	Assume that 9^k-1 is divisible by 8.
3. Inductive step:	Show that 9^{k+1}-1 is divisible by 8.

6.

Table 8.44:

	9^k-1 divisible by 8 $\Rightarrow$ $8W = (9^k\text{-}1)$ for some integer W
	9^{k+1}-1=9(9^k-1)+8=9(8W)+8,which is divisible by 8

7.
8.

Table 8.45:

1. Base case:	If $n = 1$,x^k-1=x-1=(x-1)(1)
2. Inductive hypothesis:	Assume that x^k-1 is divisible by x-1
3. Inductive step:	Show that x^{k+1}-1 is divisible by x-1.

9.
10. x
11. k
12. -1 divisible by
13. x
14. -1
15. $\Rightarrow$
16.
17. P
18. (
19. x
20. -1)=(
21. x
22.
23. k
24. -1) for some polynomial
25. P
26.
27. x

28. k+1
29. -1=
30. x
31. (
32. x
33.
34. k
35. -1)+(
36. x
37. -1)=
38. Px
39. (
40. x
41. -1)+(
42. x
43. -1),which is divisible by
44. x
45. -1
46.

Table 8.46:

1. Base case:	If n = 1,1^2-1=1-1=0=2×0
2. Inductive hypothesis:	Assume that k^2-k is even
3. Inductive step:	Show that $(k+1)^2$-$(k+1)$ is even.

47.
48. If
49. k
50. 2
51. -
52. k
53. is even, then
54. k
55. 2
56. -
57. k
58. =2
59. M
60. for some integer
61. M
62. (
63. k
64. +1)
65. 2
66. -(
67. k
68. +1)=
69. k
70. 2

71. +2
72. k
73. +1-
74. k
75. -1=
76. k
77. ^{2}
78. -
79. k
80. +2
81. k
82. =2
83. M
84. +2
85. k
86. =2(
87. M
88. +
89. k
90.),which is even because
91. M
92. +
93. K
94. is an integer.
95.

Table 8.47:

1. Base case:	If n = 1,5^{1}+1=5+1=6=6(1)
2. Inductive hypothesis:	Assume that 5^{2k-1}+1 is divisible by 6.
3. Inductive step:	Show that $5^{2(k+1)-1}$+1</math> is divisible by 6.

96.
97. If 5
98. $^{2k-1}$
99. +1 is divisible by 6, then 5
100. $^{2k-1}$
101. +1 =6M for some integer
102. M
103. .
104. 5
105. $^{2(k+1)-1}$
106. }+1 = 5
107. $^{2k+1}$
108. +1= 5
109. ^{2}
110. (5
111. $^{2k-1}$
112. +1)-24=5
113. ^{2}

114. (6M)-24 which is divisible by 6.
115.

Table 8.48:

1. Base case:	If $n = 3$,$2(3)+1 = 7$, $2^3 = 8$. $7 < 8$ so the base case is true.
2. Inductive hypothesis: 3. Inductive step:	Assume that $2k+1< 2^k$ for $k > 3$ Show that $2(k+1)+1 < 2^{k+1}$

116.
117. 2(
118. k
119. +1)+1 =2
120. k
121. +2+1=(2k+1)+2<2
122. k
123. +2<2
124. k
125. +2
126. k
127. =2(2
128. k
129.)=2
130. k+1
131.

Table 8.49:

1. Base case:	If $n = 1$, $3^1 = 3$ and $1^2 =1$. $3 > 1$ so the base case is true.
2. Inductive hypothesis: 3. Inductive step:	Assume that $3^k > k^2$ Show that $3^{k+1} > (k+1)^2$

132.
133. (
134. k
135. +1)
136. 2
137. =
138. k
139. 2
140. +2
141. k
142. +1<3
143. k
144. +2
145. k
146. +1<3

147. ^{k}
148. $+2$
149. ^{k}
150. <3
151. ^{k}
152. $+3$
153. ^{k}
154. $<3 \times 3$
155. ^{k}
156. $=3$
157. $^{k+1}$

Vocabulary

Factor A factor is a number or an expression that is multiplied with other factors to create a product.

Factorial Factorial refers to the product of the positive integers from 1 to some value n:
$n!=1 \times 2 \times 3 \times 4...\times (n\text{-}1) \times n$

Inequality An inequality is a statement that two quantities are not equal.

Postulate A postulate is a statement that is accepted as true without proof.

8.5 Geometric Series

Learning objectives

- Find the value of a finite geometric series.
- Determine if a geometric series converges or diverges.
- If an infinite series converges, find the sum.

Introduction

In lesson 2 we began working with series (sums), and in lesson 4 we proved formulas for certain sums, such as the sum of terms in an arithmetic series. In this lesson we will focus on **geometric series**, the sum of terms from a geometric sequence. We will prove a formula that allows us to find the sum of a finite number of terms, and we will return to the issue of the sum of an infinite number of terms.

Finite Geometric Series

A finite geometric series is simply a geometric series with a specific number of terms. For example, consider again the series introduced in lesson 2: 50+25+12.5+....The series is geometric: the first term is 50, and the common ration is (1/2).

The sum of the first two terms is $50 + 25 = 75$. We can write this as S_2=75
The sum of the first three is $50 + 25 + 12.5 = 87.5$. We can write this as S_3=87.5

To find the value of S_n in general, we could simply add together the first n terms in a series. However, this would obviously be tedious for a large value of n. Given the regular pattern in a geometric series - every term is $(1/r)$ of the previous term, and the n^{th} term is $a_n = a_1 r^{n-1}$, we can use induction to prove a formula for S_n .

Claim: The sum of the first n terms in a geometric series is $S_n = \frac{a_1(1-r^n)}{1-r}$

Proof by induction:

1. If $n = 1$, the n^{th} sum is the first sum, or a_1 . Using the hypothesized equation, we get $S_1 = \frac{a_1(1-r^1)}{1-r} = \frac{a_1(1-r)}{1-r} = a_1$. This establishes the base case.
2. Assume that the sum of the first k terms in a geometric series is $S_k = \frac{a_1(1-r^k)}{1-r}$.
3. Show that the sum of the first k+1 terms in a geometric series is $S_{k+1} = \frac{a_1(1-r^{k+1})}{1-r}$.

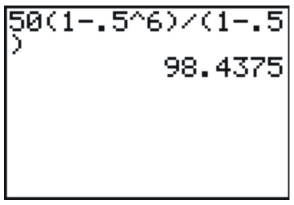
Table 8.50:

$S_{k+1} = S_k + a_{k+1}$	The $k+1$ sum is the k^{th} sum, plus the $k+1$ term
$= \frac{a_1(1-r^k)}{1-r} + a_1 r^{k+1-1}$	Substitute from step 2, and substitute the $k+1$ term
$= \frac{a_1(1-r^k)}{1-r} + \frac{a_1 r^k(1-r)}{1-r}$	The common denominator is $1 - r$
$= \frac{a_1(1-r^k)+a_1 r^k(1-r)}{1-r}$	Simplify the fraction
$= \frac{a_1\left[1-r^k+r^k(1-r)\right]}{1-r}$	
$= \frac{a_1\left[1-r^k+r^k-r^{k+1}\right]}{1-r}$	
$= \frac{a_1\left[1-r^{k+1}\right]}{1-r}$	It is proven.

Therefore we have shown that $S_n = \frac{a_1(1-r^n)}{1-r}$ for a geometric series. Now we can use this equation to find any sum of a geometric series. For example, for the series 50 + 25 + 12.5 + ... , the sum of the first 6 terms is:

$S_n = \frac{a_1(1-r^n)}{1-r} = \frac{50\left(1-(\frac{1}{2})^6\right)}{1-\frac{1}{2}} = \frac{50\left(1-\frac{1}{64}\right)}{\frac{1}{2}} = \frac{50\left(\frac{63}{64}\right)}{\frac{1}{2}} = 50\left(\frac{63}{64}\right)\left(\frac{2}{1}\right) = 98\frac{7}{16}$

The figure below shows the same calculation on a TI-83/4 calculator:



We can use this formula, as long as the series in question is geometric.

Example 1: Find the sum of the first 10 terms of a geometric series with $a_1 = 3$ and $r = 5$.

Solution: The sum is 58,593.

$S_n = \frac{a_1(1-r^n)}{1-r} = \frac{3(1-5^7)}{1-5} = \frac{3(1-78{,}125)}{-4} = \frac{3(-78{,}124)}{-4} = 58,593$

Notice that because the common ratio in this series is 5, the terms get larger and larger. This means that for increasing values of n the sums will also get larger and larger. In contrast, in the series with

common ratio (1/2), the terms gets smaller and smaller. This situation implies something important about the sum.

Infinite Geometric Series

Now let's return to the situation in the introduction in lesson 2: you are walking a 100-mile charity walk. You walk halfway, 50 miles. Then you walk half of that, 25. If you keep walking half of the remaining distance, how will you ever finish the walk?

At the end of lesson 2 we represented the sum of the infinitely many pieces of this walk as: $\sum_{n=1}^{\infty} 50\left(\frac{1}{2}\right)^{n-1}$. We know that the pieces have to add up to 100 miles. But how is it possible for the sum of an infinite number of terms to be a finite number?

To find the sum of an infinite number of terms, we should consider some partial sums. Above we found three partial sums: $S_2 = 75, S_3 = 87.5$, and $S_6 = 98\frac{7}{16}$ or 98.4375. Now let's look at larger values of n :

Table 8.51:

S_7	$= \frac{50\left(1-\left(\frac{1}{2}\right)^7\right)}{1-\frac{1}{2}} \approx 99.2$
S_8	$= \frac{50\left(1-\left(\frac{1}{2}\right)^8\right)}{1-\frac{1}{2}} \approx 99.6$
S_{10}	$= \frac{50\left(1-\left(\frac{1}{2}\right)^{10}\right)}{1-\frac{1}{2}} \approx 99.9$

As n approaches infinity, the value of S_n seems to approach 100. In terms of the actual sums, what is happening is this: as n increases, the n^{th} term gets smaller and smaller, and so the n^{th} term contributes less and less to the value of S_{n} . We say that the series **converges,** and we can write this with a limit:

Table 8.52:

$\lim_{n\to\infty} S_n$	$= \lim_{n\to\infty}\left(\frac{50\left(1-\left(\frac{1}{2}\right)^n\right)}{1-\frac{1}{2}}\right)$
	$= \lim_{n\to\infty}\left(\frac{50\left(1-\left(\frac{1}{2}\right)^n\right)}{\frac{1}{2}}\right)$
	$= \lim_{n\to\infty}\left(100\left(1-\left(\frac{1}{2}\right)^n\right)\right)$

As n approaches infinity, the value of $\left(\frac{1}{2}\right)^n$ gets smaller and smaller. That is, the value of this expression approaches 0. Therefore the value of $1-\left(\frac{1}{2}\right)^n$ approaches 1, and $100\left(1-\left(\frac{1}{2}\right)^n\right)$ approaches $100(1) = 100$.

We can do the same analysis for the general case of a geometric series, as long as the terms are getting smaller and smaller. This means that the common ratio must be a number between -1 and 1: $|\text{r}| < 1$.

Table 8.53:

$\lim_{n\to\infty} S_n$	$= \lim_{n\to\infty}\left(\frac{a_1(1-r^n)}{1-r}\right)$
	$= \frac{a_1}{1-r}$, as $(1-r^n) \to 1$

Therefore, we can find the sum of an infinite geometric series using the formula $S = \frac{a_1}{1-r}$.

Example 2: Determine if the series converges. If it converges, find the sum.

Table 8.54:

a. $40 + -20 + 10 + -5 + ...$	b. $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + ...$	c. $3 + -6 + 12 + -24 + ...$

Solution:

a. $40 + -20 + 10 + -5 + ...$ converges
The common ratio is $\frac{-1}{2}$. Therefore the sum converges to:

Table 8.55:

$\frac{40}{1-\left(\frac{-1}{2}\right)} = \frac{40}{\frac{3}{2}} = 40\left(\frac{2}{3}\right) = \frac{80}{3}$

b. $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + ...$ converges.
The common ratio is (1/3) . Therefore the sum converges to:

Table 8.56:

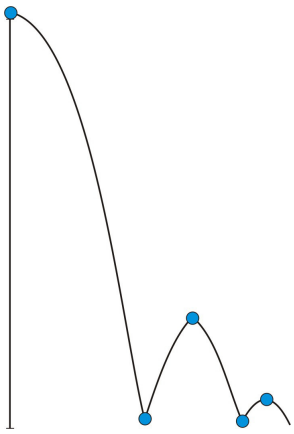
$\frac{1}{1-\frac{1}{3}} = \frac{1}{\frac{2}{3}} = \frac{3}{2}$

c. The series 3 + -6 + 12 + -24 + ... does not converge, as the common ratio is -2.

Remember that the idea of an infinite sum was introduced in the context of a realistic situation, albeit a paradoxical one. We can in fact use infinite geometric series to model other realistic situations. Here we will look at another example: the total vertical distance traveled by a bouncing ball.

The Infinitely Bouncing Ball

Consider a bouncing ball that bounces to a particular percent or fraction of its previous height with each bounce. If the ball is left alone, theoretically, it will continue this pattern infinitely.



We can think of the total distance as the distance the ball travels down plus the distance the ball travels back up. The downward bounces form a geometric series, and the upward bounces form a geometric series. So we can think of the total distance as the sum of two series. (You can also develop one series, which you will do in the review questions). The example below shows how to find the total distance. **Example 3:** A bouncing ball

A ball is dropped from a height of 20 feet. Each time it bounces, it reaches 50% of its previous height. What is the total vertical distance the ball travels?

Solution:

We can think of the total distance as the distance the ball travels down + the distance the ball travels back up. The downward bounces form a geometric series:

$20 + 10 + 5 + ...$

The upward bounces form the same series, except the first term is 10.
So the total distance is: $\sum_{n=1}^{\infty} 20\left(\frac{1}{2}\right)^{n-1} + \sum_{n=1}^{\infty} 10\left(\frac{1}{2}\right)^{n-1}$.
Each sum converges, as the common ratio is (1/2). Therefore the total distance is:

Table 8.57:

$\frac{20}{1-\frac{1}{2}} + \frac{10}{1-\frac{1}{2}} = \frac{20}{\frac{1}{2}} + \frac{10}{\frac{1}{2}} = 40 + 20 = 60$

So the ball travels a total vertical distance of 60 feet.

Lesson Summary

In this lesson we have found the sums of finite geometric series, and we have returned to the idea of an infinite series. A geometric series is either convergent or divergent. In this lesson you have learned to determine if a geometric series converges or diverges: a geometric series will converge if $|r| < 1$, and the sum is $S = \frac{a_1}{1-r}$. It is important to note that other kinds of series converge besides geometric series. You will examine these other types of series when you study calculus.

Points to Consider

1. How can infinitely many numbers add up to a finite number?
2. How does induction help us prove formulas that simplify large sums?

Review Questions

1. Find the sum of each series: a. The first term of a geometric series is 4, and the common ratio is 3. Find S_8.
   b. The first term of a geometric series is 80, and the common ratio is (1/4). Find S_7.
2. Find the sum: 5+10+20+...+640 (Hint: if a_n=640 , what is n?)
3. Use a geometric series to answer the question: In January, a company's sales totaled $11,000. It is predicted that the company's sales will increase 5% each month for the next year. At this rate, what will be the total sales for the year?
4. For questions 4 – 6, determine if the series converges or diverges. If it converges, find the sum.
5. -3+6+-12+24+...
6. 240+60+15+...
7. 9+6+4+(8/3)+...
8. In this lesson, we proved the formula for the sum of a geometric series, $S_n = \frac{a_1(1-r^n)}{1-r}$, using induction. Here you will prove this formula without induction: (1) Let S_n=a_1+$a_1 r$+$a_1 r^2$+...+$a_1 r^{n-1}$
   (2) Multiply S_n by r to obtain a second equation
   (3) Subtract the equations and solve for S_n.
9. A ball is dropped from a height of 40 feet, and each time it bounces, it reaches 25% of its previous height. a. Find the total vertical distance the ball travels, using the method used in the lesson.
   b. Find the total vertical distance the ball travels using a single series.
   (Hint: write out several terms for each bounce. For example, the first bounce is: 40 feet down + 10 feet up = 50 feet traveled.)
10. Below are two infinite series that are not geometric. Use a graphing calculator to examine partial sums. Does either series converge? a. $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ...$
    b. $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + ...$

Review Answer

1. a.$S_8 = \frac{4(1-3^8)}{1-3} = 13,120$b.$S_7 = \frac{80\left(1-\left(\frac{1}{4}\right)^7\right)}{1-\frac{1}{4}} \approx 106.66$
2. $S_8 = \frac{5(1-2^8)}{1-2} = 1275$
3. $S_{12} = \frac{11000(1-1.05^{12})}{1-1.05}$ =175,088.39 $The sum does not converge because r = -2.$
4. The sum converges. $S = 320$.
5. The sum converges. $S = 27$.
6. Follow the steps below:

Table 8.58:

(1) $S_n = a_1+a_1r+a_1r^2+...+a_1r^{n-1}$
(2) $rS_n = a_1r + a_1r^2 + a_1r^3 + ... + a_1r^n$ (3) $S_n - rS_n = a_1 - a_1r^n$

Table 8.58: (continued)

	(1) $S_n = a_1 + a_1 r + a_1 r^2 + \ldots + a_1 r^{n-1}$
	$\Rightarrow S_n(1-r) = a(1-r^n)$ $\Rightarrow S_n = \frac{a(1-r^n)}{(1-r)}$

7.
8. a.$\sum_{n-1}^{\infty} 40\left(\frac{1}{4}\right)^{n-1} + \sum_{n-1}^{\infty} 20\left(\frac{1}{4}\right)^{n-1} = 66\frac{2}{3}$b.$\sum_{n-1}^{\infty} 50\left(\frac{1}{4}\right)^{n-1} = 66\frac{2}{3}$
9. a.This series does not converge. b.This series converges around 1.65. (The actual sum is $\frac{\pi^2}{6}$.)

Vocabulary

Converge If the limit of the partial sums of series exists and is finite, the series converges.

Diverge If the limit of the partial sums of series does not exist or is infinite, then the series diverges.

Geometric series A geometric series is the sum of terms from a geometric sequence.

8.6 The Binomial Theorem

Learning objectives

- Work with combinations.
- Generate rows of Pascal's triangle.
- Use the binomial theorem to expand polynomials, and to identify terms for a given polynomial.
- Use the binomial theorem to calculate the probability of success or failure in a Bernoulli trial.

Introduction

In this lesson we will develop and prove the Binomial Theorem. This theorem tells us about the relationship between two forms of a certain type of polynomial: the factored, binomial form $(x + y)^n$, and the expanded form of this polynomial. To understand the theorem, we will first revisit a topic you have seen previously, factorials, and we will discuss several other ideas needed to develop the theorem. Then we will prove the theorem using induction. Finally, we will use the theorem to expand polynomials and to solve a certain type of probability problem.

Preliminary ideas: factorials, combinations, and Pascal's Triangle

Recall that a factorial of a positive integer n is the product of n, and all of the positive integers less than n. We write this as $n! = n(n - 1)(n - 2) \ldots (3)\ (2)\ (1)$.

In order to develop the binomial theorem, we need to look at a related idea: combinations. If you have studied probability, you may be familiar with combinations and permutations. A **combination** is the number of ways you can choose r objects from a group of n objects, if the order of choosing does not matter. An example will help clarify the idea of a combination:

In a class of 20 students, 3 students are going to be chosen to form a committee to plan a fieldtrip. How many possible committees are there?

To answer this question, we need to figure out how many ways we can choose groups of 3 students from the 20 on the class. The order of choosing does not matter. That is, if I choose Amy, Juan, and Nina, it is the same as choosing Juan, then Amy, then Nina, or any other ordering of the three students.

In general, we can find the number of combinations of r objects chosen from n objects by the following:

Table 8.59:

${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

(Note that there are two different symbols for combinations: ${}_nC_r$ and $\binom{n}{r}$ You can use either one, though ${}_nC_r$ is what is used on the TI-83/84.)

Therefore the number of combinations of 3 people from 20 people in the class is

Table 8.60:

$\frac{20!}{3!17!} = 1140$

A note about calculating combinations: for small numbers, it is not too difficult to calculate by hand. For example, $\binom{7}{4} = \frac{7!}{4!3!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 7 \cdot 5 = 35$.

Canceling factors in the numerator and denominator simplifies the calculation. However, you can always use a graphing calculator. For example, to find $\binom{20}{3}$, press: 20 <TI font_MATH> and then move right to the **PRB** menu. Press 3. This takes you back to the main screen. You should see 20 ${}_nC_r$. Now press 3 >,TI font_ENTER>. You should see the answer, 1140.

There is a particular pattern in combinations that is seen in the expansion of polynomials of the form $(x + y)^n$. This pattern is most commonly displayed in a triangle:

Table 8.61:

$\binom{0}{0}$

Table 8.62:

$\binom{1}{0}\binom{1}{1}$

Table 8.63:

$\binom{2}{0}\binom{2}{1}\binom{2}{2}$

Table 8.64:

$\binom{3}{0}\binom{3}{1}\binom{3}{2}\binom{3}{3}$

Table 8.65:

$\binom{4}{0}\binom{4}{1}\binom{4}{2}\binom{4}{3}\binom{4}{4}$

Table 8.66:

$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$

If we evaluate the combinations in the triangle, we get:

1

1 1

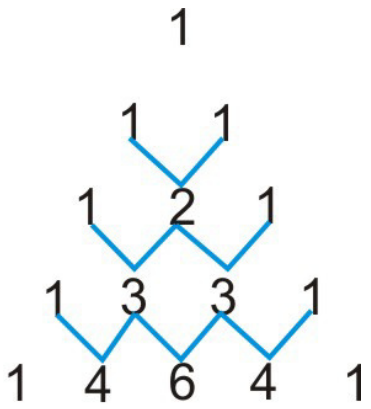
1 2 1

1 3 3 1

1 4 6 4 1

etc.

This triangle is referred to as Pascal's triangle, named after mathematician Blaise Pascal, although other mathematicians before him worked with these numbers. The numbers in the triangle can be used to generate more rows: notice that if you add two consecutive numbers, you get the number between and below them in the next row:



etc.

We can generalize this pattern as follows: $\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$.

We can prove this statement directly:

Table 8.67:

$\binom{n}{r-1} + \binom{n}{r} = \frac{n!}{(r-1)!(n-(r-1))!} + \frac{n!}{r!(n-r)!}$

Table 8.68:

$= \frac{n!}{(r-1)!(n-r+1))!} + \frac{n!}{r!(n-r)!}$	
$= \frac{rn!}{r(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r)!}$	Multiply the first fraction by $\frac{r}{r}$. Substitute $r(r-1)!$ with $r!$
$= \frac{rn!}{r!(n-r+1)!} + \frac{n!(n-r+1)}{r!(n-r)!(n-r+1)}$	Multiply the second fraction by $\frac{(n-r+1)}{(n-r+1)}$. Substitute $(n-r+1)(n-r)!$ with $(n-r+1)!$
$= \frac{rn!}{r!(n-r+1)!} + \frac{n!(n-r+1)}{r!(n-r+1)!}$	Now the fractions have a common denominator. Add the fractions and simplify the numerator:
$= \frac{rn!+n!(n-r+1)}{r!(n-r+1)!}$	
$= \frac{n!(r+n-r+1)}{r!(n-r+1)!} = \frac{n!(n+1)}{r!(n-r+1)!}$	Factor out $n!$. Simplify $r+n-r+1 = n+1$
$= \frac{(n+1)!}{r!(n-r+1)!} = \frac{(n+1)!}{r!(n+1-r)!} = \binom{n+1}{r}$.	

This proves that we can always generate rows of Pascal's triangle. Further, the fact that $\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$ will be used below to prove the Binomial Theorem.

Binomial Expansions

To expand a binomial is to multiply all of the factors. The resulting polynomial is in standard form. For example:

$(x + y)^2 = (x + y)\ (x + y) = x^2 + xy + xy + y^2 = x^2 + 2xy + y^2$

If we expand $(x + y)^3$, we get:

Table 8.69:

$(x + y)^3$	$= (x + y)\ (x + y)\ (x + y)$
	$= (x + y)\ (x^2 + 2xy + y^2)$ $= x^3 + 2x^2y + xy^2 + x^2y + 2xy^2 + y^3$ $= x^3 + 3x^2y + 3xy^2 + y^2$

Notice that the coefficients of each polynomial correspond to a row of Pascal's triangle. Also notice that the exponents of x descend, and the exponents of y ascend with each term. These are key aspects of the

The Binomial Theorem

The binomial theorem can be stated using a summation:

Table 8.70:

$(x+y)^n = \sum_{r=0}^{n} \left(\binom{n}{r} x^{n-r} y^r \right)$

This is a very succinct way of summarizing the pattern in a binomial expansion. Let's return to $(x + y)^3$ to see how the theorem works.

Table 8.71:

$(x+y)^3$	$= \binom{3}{0}x^{3-0}y^0 + \binom{3}{1}x^{3-1}y^1 + \binom{3}{2}x^{3-2}y^2 + \binom{3}{3}x^{3-3}y^3$
	$= 1x^3 \cdot 1 + 3x^2y + 3xy^2 + 1 \cdot x^0 \cdot y^3$ $= x^3 + 3x^2y + 3xy^2 + y^3$

Again, the exponents on x descend from 3 to 0. The exponents on y ascend from 0 to 3. The coefficients on the terms correspond to row 3 of Pascal's triangle. These coefficients are, not surprisingly, referred to as binomial coefficients!

Given this theorem, we can expand any binomial without having to multiply all of the factors. We can also identify a particular term in an expansion. Before we solve these kinds of problems, we will prove the theorem:

The Binomial Theorem: $(x+y)^n = \sum_{r=0}^{n} \left(\binom{n}{r} x^{n-r} y^r \right)$

Proof by induction:

1. The base case
   If $n = 1$, then the binomial is $(x + y)^1 = x + y = 1x + 1y$. The coefficients correspond to row 1 (physically the second row) of the triangle.
   Note: we have already observed that the theorem holds for $n = 2$ and $n = 3$.
2. The inductive hypothesis:
   Assume that $(x+y)^k = \sum_{r=0}^{k} \left(\binom{k}{r} x^{k-r} y^r \right)$.
3. The inductive step: show that $(x+y)^{k+1} = \sum_{r=0}^{k+1} \left(\binom{k+1}{r} x^{k+1-r} y^r \right)$

Table 8.72:

$(x+y)^{k+1} = (x+y)(x+y)^k$

Table 8.73:

$= (x+y)(x^k + \sum_{r=1}^{k-1} \binom{k}{r} x^{k-r}y^r + y^k)$	Use the inductive hypothesis, but write out the first and last term: x^k, y^k
$= x^{k+1} + y^{k+1} + yx^k + xy^k + x\left(\sum_{r=1}^{k-1} \binom{k}{r} x^{k-r}y^r\right) + y\left(\sum_{r=1}^{k-1} \binom{k}{r} x^{k-r}y^r\right)$	Multiply
$= x^{k+1} + y^{k+1} + yx^k + xy^k + \sum_{r=1}^{k-1} \binom{k}{r} x^{k+1-r}y^r + \sum_{r=1}^{k-1} \binom{k}{r} x^{k-r}y^{r+1}$	yx^k is $\binom{k}{0}x^{k-0}y^{0+1}$ and
$= x^{k+1} + y^{k+1} + \sum_{r=1}^{k-1} \binom{k}{r} x^{k+1-r}y^r + \sum_{r=0}^{k-1} \binom{k}{r} x^{k-r}y^{r+1}$	xy^k is $\binom{k}{1}x^{k+1-1}y^1$
$= x^{k+1} + y^{k+1} + \sum_{r=1}^{k-1} \binom{k}{r} x^{k+1-r}y^r + \sum_{j=1}^{k} \binom{k}{j-1} x^{k+1-j}y^j$	Let j = r+1 and rewrite the sum.
$= x^{k+1} + y^{k+1} + \sum_{r=1}^{k-1} \binom{k}{r} x^{k+1-r}y^r + \sum_{r=1}^{k} \binom{k}{r-1} x^{k+1-r}y^r$	Then rewrite the sum again with r.
$= x^{k+1} + y^{k+1} + \sum_{r=1}^{k} \left[\binom{k}{r} + \binom{k}{r-1}\right] x^{k+1-r}y^r$	It was proven above: $\binom{n}{r-1} + nr - 1 = \binom{n+1}{r}$
$= x^{k+1} + y^{k+1} + \sum_{r=1}^{k} \binom{k+1}{r} x^{k+1-r}y^r + \sum_{r=0}^{k+1} \binom{k+1}{r-1} x^{k+1-r}y^r$	

This proves the Binomial Theorem. Now we can use the theorem to expand binomials, and to identify terms and coefficients in an expansion.

Expansion, and Terms and Coefficients in a Binomial Expansion

As noted above, the binomial theorem allows us to expand binomials. So far we have only seen the binomial $(x + y)^n$, but the two terms in the binomial could be any monomials.

Example 1: Use the binomial theorem to expand each polynomial:

Table 8.74:

a.$(2x + a)^4$	b. $(x - 3)^5$

Solution:

a.$(2x + a)^4$

Table 8.75:

	$= 1(2x)^4 (a)^0 + 4(2x)^3 (a)^1 + 6(2x)^2(a)^2 + 1(2x)^0 (a)^4$
	$= 16x^4 + 32x^3a + 24x^2a^2 + 8x'a^3 + a^4$

Note that it is easier to simply use the numbers from the appropriate row of the triangle than to write out all of the coefficients as combinations. However, if n is large, it may be easier to use the

combinations.

b. $(x - 3)^5$

Table 8.76:

	$= 1(x)^5 (-3)^0 + 5(x)^4 (-3)^1 + 10(x)^3(-3)^2 + 10(x)^2(-3)^3 + 5(x)^1 (-3)^4 + 1(x)^0 (-3)^5$
	$= x^5 - 1x^4a + 90x^3 - 270x^2 + 405x - 243$

Notice that in this expansion, the terms alternate signs. This is the case because the second term in the binomial is -3. When expanding this kind of polynomial, be careful with your negatives!

We can also use the Binomial Theorem to identify a particular term or coefficient.

Example 2: Identify the 3rd term of the expansion of $(2x + 3)^6$.

Solution: The 3rd term is $\binom{6}{2}(2x)^4 3^2 = 15 \cdot 16x^4 \cdot 9 = 2160x^4$

Keep in mind that row 6 of Pascal's triangle starts with $\binom{6}{0}$, so the coefficient of the third term in the expansion is $\binom{6}{2}$. Also keep in mind that the first term includes $(2x)^6$ and 3^0, so the third term includes $(2x)^4$ and 3^2.

Finding a term in an expansion can be used to answer a particular kind of probability question. Consider an experiment, in which there are two possible outcomes, such as flipping a coin. If we flip a coin over and over again, this is referred to as a **Bernoulli trial**. In each flip ("experiment"), the probability of getting a head is 0.5, and the probability of getting a tail is 0.5. (Note: this is true for flipping a coin, but not for other situations. That is, it's not always "50-50 chance!) Now say we flip a coin 25 times. What is the probability of getting exactly 10 heads?

The answer to this question is a term of a binomial expansion. That is, the probability of getting 10 heads from 25 coin tosses is: $\binom{25}{10}(0.5)^{10}(0.5)^{15} \approx 0.0974$, or about a 10% chance. Now let's consider an example that is not a "50-50" situation.

Example 3: At the carnival, you decide to play a game of chance. You buy 15 tickets for the game. You have a 75% chance of winning each time you play the game. What is the probability that you will win exactly 8 of the 15 games?

Solution: About 4%

For each of the 15 games, there is 75% chance of winning, and a 25% chance of losing. So the probability of exactly 8 wins is $\binom{15}{8}(.75)^8(.25)^7 \approx 0.039$, or about a 4% chance. It is important to note that this is only the probability of winning exactly 8 games. If for example, you need to win 8 or more games in order to win a certain prize, the probability of winning would be higher, as it would be the sum of the probabilities of winning 8, 9, 10, 11, 12, 13, 14, and 15 games.

Summary

In this lesson we have developed, proved, and used the Binomial Theorem. This theorem is useful because it allows us to expand polynomials without having to do many steps of multiplication of factors. It is also useful because it allows us to identify a term of an expansion, which can be used to solve certain

probability problems. In your study of calculus, you will also find that binomial expansions are useful for approximating certain polynomials.

Points to Consider

1. Pascal's triangle is named after mathematician Blaise Pascal, even though he was not the first person to work with these numerical relationships. Why might this be the case?
2. What happens to a binomial expansion if n is not a positive integer?
3. Why is the probability of winning k out of n games not k/n?

Review Questions

1. Write out the combination and find its value: a. ${}_8C_5$
   b. ${}_6C_3$
2. In a class of 200 students, 25 will be chosen randomly to participate in a research study. How many possible groups of 25 students can be chosen? (Hint: use a calculator!)
3. Expand: $(x+3a)^4 =$
4. Expand: $(y+\frac{1}{2})^5 =$
5. Expand: $(2x-a)^6 =$
6. Find the 3^{rd} term in the expansion $(3x+2a)^9$.
7. Find the 7^{th} term in the expansion of $(4x-\frac{1}{2}a)^{10}$.
8. A die is rolled 10 times. What is the probability of rolling exactly four 4's? (Hint: the probability of rolling a 4 is 1/6. The probability of not rolling a 4 is 5/6.)
9. The local TV station forecasts a 30% chance of rain every day for the next week. What is the probability that it will rain on exactly 6 out of the next 7 days?
10. Consider the following situation: a basketball player is going to attempt to make 20 free throws. She is assuming that she has an 80% chance of making each shot. What is the probability that she will make exactly 19 out of 20 shots? Two students were discussing this problem. Student A says that the problem is a Bernoulli trial, and that the probability is $\binom{20}{19}(.8)^{19}(.2)^1$. Student B disagrees, and says that the situation is *not* a Bernoulli trial. What reasoning might student B use to support his argument?

Review Answers

1. a. $\frac{8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{5\cdot4\cdot3\cdot2\cdot1\cdot3\cdot2\cdot1} = \frac{8\cdot7\cdot6}{3\cdot2\cdot1} = 8\cdot7 = 56$ b. $\frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot3\cdot2\cdot1} = \frac{6\cdot5\cdot5}{3\cdot2\cdot1} = 5\cdot4 = 20$
2. ${}_{200}C_{25} \approx 4.52\times10^{11}$
3. $x^4 + 12x^3a + 54x^2a^2 + 108xa^3 + 81a^4$
4. $y^5 + \frac{5}{2}y^4 + \frac{5}{2}y^3 + \frac{5}{4}y^2 + \frac{5}{16}y + \frac{1}{32}$
5. $64x^6 + 19.2x^5 + 2.4x^4 + 0.16x^3 + 0.006x^2 + 0.00012x^5 + 0.000001x^6$
6. $\binom{9}{3}(3x)^7(2a)^2 = 8.748x^7a^4$
7. $\binom{10}{7}(4x)^4(-\frac{1}{6}a)^6 = \frac{160}{243}x^4a^6$
8. About 5.4%
9. About 0.36%
10. If the situation is a Bernoulli trial, then the probabilities of success and failure are the same with each shot. The player has assumed this, but this might not be the case. For example, if she improves as she shoots, the probability of success increases. Or, if she gets tired, distracted, or demotivated, the probability of success might decrease.

Vocabulary

Bernoulli trial A Bernoulli trial is a series of experiments in which each experiment has two possible outcomes, usually referred to as "success" and "failure."

Binomial A binomial is a polynomial with two terms.

Combination A combination refers to choosing one or more objects from a set when the order of choosing does not matter.

Factorial Factorial refers to the product of the positive integers from 1 to some value n:$1 \times 2 \times 3 \times 4....\times(n\text{-}1)\times n = n!$

Monomial A monomial is a term of the form ax^{n} , where a is a real number and n is a whole number.

Polynomial A polynomial is a sum of monomials.